Master of Business Administration

(Open and Distance Learning Mode)

Semester – II

Business Research Methods

Centre for Distance and Online Education (CDOE)

DEVI AHILYA VISHWAVIDYALAYA, INDORE

"A+" Grade Accredited by NAAC

IET Campus, Khandwa Road, Indore - 452001

www.cdoedavv.ac.in

www.dde.dauniv.ac.in

CDOE-DAVV

Program Coordinator

Dr. Manishkant Arya
Centre for Distance and Online Education (CDOE)
Devi Ahilya Vishwavidyalaya, Indore – 452001

Content Design Committee

Dr. Sangita Jain
Institute of Management Studies
Devi Ahilya Vishwavidyalaya, Indore – 452001

Dr. Yamini Karmarkar
Institute of Management Studies
Devi Ahilya Vishwavidyalaya, Indore – 452001

Dr. Geeta Neema
International Institute of Professional Studies
Devi Ahilya Vishwavidyalaya, Indore – 452001

Dr. Manishkant Arya
Centre for Distance and Online Education (CDOE)
Devi Ahilya Vishwavidyalaya, Indore - 452001

Language Editors

Dr. Arti Sharan
Institute of Engineering & Technology
Devi Ahilya Vishwavidyalaya, Indore – 452001

Dr. Ruchi Singh
Institute of Engineering & Technology
Devi Ahilya Vishwavidyalaya, Indore – 452001

SLM Author(s)

Mr. Sandeep Handa
MBA
EMRC, Devi Ahilya Vishwavidyalaya, Indore – 452001

Mr. Aditya Singh Parmar
MBA
EMRC, Devi Ahilya Vishwavidyalaya, Indore – 452001

Business Research Methods

MBA

Table of Contents

UNIT 4 SAMPLING

UNIT 5 MEASUREMENT AND SCALING TECHNIQUES

UNIT 1 INTRODUCTION TO BUSINESS RESEARCH

STRUCTURE

1.0 OBJECTIVES

After studying this unit, you should be able to:

l explain the meaning of research,

l differentiate between Science and Knowledge,

l distinguish between inductive and deductive logic,

l discuss the need for research in business,

l classify research into different types,

l narrate different methods of research,

l list the difficulties in business research, and

l explain the business research process and its role in decision making.

1.1 INTRODUCTION

Research is a part of any systematic knowledge. It has occupied the realm of human understanding in some form or the other from times immemorial. The thirst for new areas of knowledge and the human urge for solutions to the problems, has developed a faculty for search and research and re-research in him/her. Research has now become an integral part of all the areas of human activity.

Research in common parlance refers to a search for knowledge. It is an endeavour to discover answers to problems (of intellectual and practical nature) through the application of scientific methods. Research, thus, is essentially a systematic inquiry seeking facts (truths) through objective, verifiable methods in order to discover the relationship among them and to deduce from them broad conclusions. It is thus a method of critical thinking. It is imperative that any type of organisation in the globalised environment needs systematic supply of information coupled with tools of analysis for making sound decisions which involve minimum risk. In this Unit, we will discuss at length the need and significance of research, types and methods of research, and the research process.

1.2 MEANING OF RESEARCH

The Random House Dictionary of the English language defines the term 'Research' as a diligent and systematic inquiry or investigation into a subject in order to discover or revise facts, theories, applications, etc. This definiton explains that research involves acquisition of knowledge. Research means search for truth. Truth means the quality of being in agreement with reality or facts. It also means an established or verified fact. To do research is to get nearer to truth, to understand the reality. Research is the pursuit of truth with the help of study, observation, comparison and experimentation. In other words, the search for knowledge through objective and systematic method of finding solution to a problem/answer to a question is research. There is no guarantee that the researcher will always come out with a solution or answer. Even then, to put it in Karl Pearson's words "there is no short cut to truth… no way to gain knowledge of the universe except through the gate way of scientific method". Let us see some definitions of Research:

L.V. Redman and A.V.H. Mory in their book on "The Romance of Research" defined research as "a systematized effort to gain new knowledge"

"Research is a scientific and systematic search for pertinent information on a specific topic" (C.R. Kothari, Research Methodology - Methods and Techniques)

"A careful investigation or inquiry specially through search for new facts in any branch of knowledge" (Advanced learners Dictionary of current English)

Research refers to a process of enunciating the problem, formulating a hypothesis, collecting the facts or data, analyzing the same, and reaching certain conclusions either in the form of solution to the problem enunciated or in certain generalizations for some theoretical formulation.

D. Slesinger and M. Stephenson in the Encyclopedia of Social Sciences defined research as: "Manipulation of things, concepts or symbols for the purpose of generalizing and to extend, correct or verify knowledge, whether that knowledge aids in the construction of a theory or in the practice of an art".

To understand the term 'research' clearly and comprehensively let us analyze the above definition.

i) Research is manipulation of things, concepts or symbols

l   manipulation means purposeful handling,

l   things means objects like balls, rats, vaccine,

l   concepts mean the terms designating the things and their perceptions about which science tries to make sense. Examples: velocity, acceleration, wealth, income.

l   Symbols may be signs indicating $+, -, \div, \times, \overline{x}, \sigma, \Sigma$, etc.

l   Manipulation of a ball or vaccine means when the ball is kept on different degrees of incline how and at what speed does it move? When the vaccine is used, not used, used with different gaps, used in different quantities (doses) what are the effects?

ii) Manipulation is for the purpose of generalizing

The purpose of research is to arrive at generalization i.e., to arrive at statements of generality, so that prediction becomes easy. Generalization or conclusion of an enquiry tells us to expect some thing in a class of things under a class of conditions.

Examples: Debt repayment capacity of farmers will be decreased during drought years.

When price increases demand falls.

Advertisement has a favourable impact on sales.

iii) The purpose of research (or generalization) is to extend, correct or verify knowledge

Generalization has in turn certain effects on the established corpus or body of knowledge. It may extend or enlarge the boundaries of existing knowledge by removing inconsistencies if any. It may correct the existing knowledge by pointing out errors if any. It may invalidate or discard the existing knowledge which is also no small achievement. It may verify and confirm the existing knowledge which also gives added strength to the existing knowledge. It may also point out the gaps in the existing corpus of knowledge requiring attempts to bridge these gaps.

iv) This knowledge may be used for construction of a theory or practice of an art

The extended, corrected or verified knowledge has two possible uses to which persons may put it.

a)   may be used for theory building so as to form a more abstract conceptual system. Eg. Theory of relativity, theory of full employment, theory of wage.

b)   may be used for some practical or utilitarian goal.

Eg. 'Salesmanship and advertisement increase sales' is the generalization. From this, if sales have to be increased, use salesmanship and advertisement for increasing sales.

Theory and practice are not two independent things. They are interdependent. Theory gives quality and effectiveness to practice. Practice in turn may enlarge or correct or confirm or even reject theory.

1.3   MEANING OF SCIENCE

The development of Science can be considered as a constant inter play between theory and facts. The word "Science" comes from the Latin word "Scientia" which means "knowledge". As we have seen earlier, research involves acquisition of knowledge. Thus Science and research are related and go hand in hand.

At one time the word science was used to denote all systematic studies or organized bodies of knowledge. Let us see some definitions.

– "Science means a branch of (accumulated) knowledge". In this sense it refers to a particular field or branch of knowledge such as Physics, Chemistry, Economics.

– "The systematized knowledge about things or events in nature is called Science".

– "Science is popularly defined as an accumulation of systematic knowledge" (Goode & Hatt).

In these definitions the words 'systematic' and 'knowledge' are very important. Knowledge refers to the goal of science, while 'systematic' refers to the 'method' that is used to reach that goal. Now a days the stress is on the method rather than the knowledge. See the following definitions:

– Knowledge not of things but of their relations.

– Science is a process which makes knowledge.

– It is the approach rather than the content that is the test of science.

– Science is a way of investigation.

– Science is a way of looking at the World.

– " The unity of all sciences consists alone in its methods, not in its material" - (Karl Pearson).

From the above definitions two broad views emerge. They are: (a) Science as organized or accumulated knowledge. (b) Science as a method / process leading to knowledge. (a) is a STATIC view where as (b) is a DYNAMIC View. The view that Science is a method rather than a field of specific subject matter is more popular.

1.4 KNOWLEDGE AND SCIENCE

Knowledge has some thing to do with knowing. Knowing may be through acquaintance or through the description of the characteristics of certain things. The things with which we can be acquainted are the things of which we are directly aware. Direct awareness may come through perception and sensation. Most of our knowledge of things is by description.

Knowing has an external reference, which may be called a fact. A fact is any thing that exists or can be conceived of. A fact is neither true nor false. It is what it is. What we claim to know is belief or judgement. But every belief cannot, however, be equated with knowledge, because some of our beliefs, even the true ones, may turn out to be false on verification. Knowledge, therefore, is a matter of degree. However, knowledge need not always be private or individual. Private knowledge may be transformed into public knowledge by the application of certain scientific and common sense procedures.

Human knowledge takes the form of beliefs or judgement about a particular phenomenon. Some beliefs may be supported by evidence and some may not. The evidence may be based on our perceptions and experiences. The beliefs which are supported by evidence are called justified beliefs. Only justified beliefs are knowledge. Ordinary belief (not supported by evidence) is not knowledge.

We have shown that knowledge requires explanations and these come in Science. Knowledge and Science are not necessarily synonymous. Science implies knowledge, but the converse is not true. Therefore, we can say that "all Sciences are knowledge, but all knowledge is not science". Scientific knowledge is unified, organized and systematic, while ordinary knowledge is a jumble of isolated and disconnected facts. Science applies special means and methods to render knowledge true and exact, but ordinary knowledge rests on observations which are not methodical. But scientific knowledge and ordinary knowledge are not different in kind, but only in degree. Scientific knowledge is more specialized, exact and organized than ordinary knowledge.

Self Assessment Exercise A

1) What do you understand by Research?

   ....
   ....
   ....

2) What is the relation between Science and Research?

   ....
   ....
   ....

3) Distinguish between Knowledge and Science.

   ....
   ....
   ....

4) What is a fact?

   ....
   ....
   ....

1.5 INDUCTIVE AND DEDUCTIVE LOGIC

A rational man does not accept any statement without empirical verification or logic. After the data / facts have been collected, processed, analyzed, we have to draw broad conclusions / generalizations. Research provides an analytical framework for the subject matter of investigation. It establishes the relationship between the different variables. The cause and effect relationship between the different variables can also be identified, leading to valuable observations, generalizations and conclusions. Inductions and deductions are also possible in systematic research.

Induction is the process of reasoning whereby we arrive at generalizations from particular facts. It is a movement of knowledge from particular observations / instances to a general rule or principle. Induction involves a passage from observed to unobserved. It involves two processes - observation and generalization. For example, if it is observed in a number of cases that when price increases less is purchased. Therefore, the generalization is " when price increases demand falls".

Deduction, on the other hand, is a way of making a particular inference from a generalization. Deduction is a movement of knowledge from a general rule to a particular case. For example, 'All men are mortal' is a general rule. Ranjit is a man. Therefore, from the general rule it can be deduced that Ranjit is also mortal'. Similarly, All M.Com. degree holders are eligible for Ph.D. in Commerce is a general statement. Praneeth is a M.Com. degree holder. Therefore, it can be deduced that Praneeth is eligible for Ph.D. in Commerce.

Empirical studies have a great potential, for they lead to inductions and deductions. Research enables one to develop theories and principles, on the one hand, and to arrive at generalizations on the other. Both are aids to acquisition of knowledge.

1.6   SIGNIFICANCE OF RESEARCH IN BUSINESS

Research is the process of systematic and indepth study or search for a solution to a problem or an answer to a question backed by collection, compilation, presentation, analysis and interpretation of relevant details, data and information. It is also a systematic endeavour to discover valuable facts or relationships. Research may involve careful enquiry or experimentation and result in discovery or invention. There cannot be any research which does not increase knowledge which may be useful to different people in different ways. Let us see the need for research to business organizations and their managers and how it is useful to them.

i)   Industrial and economic activities have assumed huge dimensions. The size of modern business organizations indicates that managerial and administrative decisions can affect vast quantities of capital and a large number of people. Trial and error methods are not appreciated, as mistakes can be tremendously costly. Decisions must be quick but accurate and timely and should be objective i.e. based on facts and realities. In this back drop business decisions now a days are mostly influenced by research and research findings. Thus, research helps in quick and objective decisions.

ii)   Research, being a fact-finding process, significantly influences business decisions. The business management is interested in choosing that course of action which is most effective in attaining the goals of the organization. Research not only provides facts and figures to support business decisions but also enables the business to choose one which is best.

iii) A considerable number of business problems are now given quantitative treatment with some degree of success with the help of operations research. Research into management problems may result in certain conclusions by means of logical analysis which the decision maker may use for his action or solution.

iv)   Research plays a significant role in the identification of a new project, project feasibility and project implementation.

v)   Research helps the management to discharge its managerial functions of planning, forecasting, coordinating, motivating, controlling and evaluation effectively.

vi)   Research facilitates the process of thinking, analysing, evaluating and interpreting of the business environment and of various business situations and business alternatives. So as to be helpful in the formulation of business policy and strategy.

vii) Research and Development (R & D) helps discovery and invention. Developing new products or modifying the existing products, discovering new uses, new markets etc., is a continuous process in business.

viii) The role of research in **functional areas** like production, finance, human resource management, marketing need not be over emphasized. Research not only establishes relationships between different variables in each of these functional areas, but also between these various functional areas.

ix) Research is a must in the production area. Product development, new and better ways of producing goods, invention of new technologies, cost reduction, improving product quality, work simplification, performance improvement, process improvement etc., are some of the prominent areas of research in the production area.

x) The purchase/material department uses research to frame alternative suitable policies regarding where to buy, when to buy, how much to buy, and at what price to buy.

xi) Closely linked with production function is marketing function. Market research and marketing research provide a major part of marketing information which influences the inventory level and production level. Marketing research studies include problems and opportunities in the market, product preference, sales forecasting, advertising effectiveness, product distribution, after sales service etc.,

xii) In the area of financial management, maintaining liquidity, profitability through proper funds management and assets management is essential. Optimum capital mix, matching of funds inflows and outflows, cash flow forecasting, cost control, pricing etc., require some sort of research and analysis. Financial institutions also (banking and non-banking) have found it essential to set up research division for the purpose of collecting and analysing data both for their internal purpose and for making indepth studies on economic conditions of business and people.

xiii) In the area of human resource management personnel policies have to be guided by research. An individual's motivation to work is associated with his needs and their satisfaction. An effective Human Resource Manager is one who can identify the needs of his work force and formulate personnel policies to satisfy the same so that they can be motivated to contribute their best to the attainment of organizational goals. Job design, job analysis, job assignment, scheduling work breaks etc., have to be based on investigation and analysis.

xiv) Finally, research in business is a must to continuously update its attitudes, approaches, products goals, methods, and machinery in accordance with the changing environment in which it operates.

1.7 TYPES OF RESEARCH

Research may be classified into different types for the sake of better understanding of the concept. Several bases can be adopted for the classification such as nature of data, branch of knowledge, extent of coverage, place of investigation, method employed, time frame and so on. Depending upon the BASIS adopted for the classification, research may be classified into a class or type. It is possible that a piece of research work can be classified under more than one type, hence there will be overlapping. It must be

remembered that good research uses a number of types, methods, & techniques. Hence rigid classification is impossible. The following is only an attempt to classify research into different types.

i) **According to the Branch of Knowledge**

Different Branches of knowledge may broadly be divided into two:

a) Life and physical sciences such as Botany, Zoology, Physics and Chemistry.

b) Social Sciences such as Political Science, Public Administration, Economics, Sociology, Commerce and Management.

Research in these fields is also broadly referred to as life and physical science research and social science research. Business education covers both Commerce and Management, which are part of Social sciences. Business research is a broad term which covers many areas.

```
                         Business Research
   ┌────────┬─────────┬─────────┬────────┬──────────┬─────────┬──────────┬──────────┐
Management Production Personnel Finance Accounting Marketing Business  Business
                                                             Policy    History
```
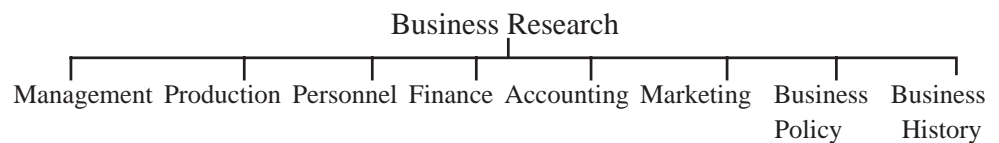
The research carried out, in these areas, is called management research, production research, personnel research, financial management research, accounting research, Marketing research etc.

Management research includes various functions of management such as planning, organizing, staffing, communicating, coordinating, motivating, controlling. Various motivational theories are the result of research. **Production (also called manufacturing) research** focuses more on materials and equipment rather than on human aspects. It covers various aspects such as new and better ways of producing goods, inventing new technologies, reducing costs, improving product quality. **Research in personnel management** may range from very simple problems to highly complex problems of all types. It is primarily concerned with the human aspects of the business such as personnel policies, job requirements, job evaluation, recruitment, selection, placement, training and development, promotion and transfer, morale and attitudes, wage and salary administration, industrial relations. Basic research in this field would be valuable as human behaviour affects organizational behaviour and productivity. **Research in Financial Management** includes financial institutions, financing instruments (egs. shares, debentures), financial markets (capital market, money market, primary market, secondary market), financial services (egs. merchant banking, discounting, factoring), financial analysis (e.g. investment analysis, ratio analysis, funds flow / cash flow analysis) etc.,

Accounting research though narrow in its scope, but is a highly significant area of business management. Accounting information is used as a basis for reports to the management, shareholders, investors, tax authorities, regulatory bodies and other interested parties. Areas for accounting research include inventory valuation, depreciation accounting, generally accepted accounting principles, accounting standards, corporate reporting etc.

Marketing research deals with product development and distribution problems, marketing institutions, marketing policies and practices, consumer behaviour, advertising and sales promotion, sales management and after sales service etc. Marketing research is one of the very popular areas and also a well established one. Marketing research includes market potentials, sales forecasting, product

testing, sales analysis, market surveys, test marketing, consumer behaviour studies, marketing information system etc.

Business policy research is basically the research with policy implications. The results of such studies are used as indices for policy formulation and implementation. **Business history research** is concerned with the past. For example, how was trade and commerce during the Moghul regime.

ii) According to the Nature of Data

A simple dichotomous classification of research is Quantitative research and Qualitative research / non-quantitative. **Quantitative research** is variables based where as qualitative research is attributes based. Quantitative research is based on measurement / quantification of the phenomenon under study. In other words, it is data based and hence more objective and more popular.

Qualitative research is based on the subjective assessment of attributes, motives, opinions, desires, preferences, behaviour etc. Research in such a situation is a function of researcher's insights and impressions.

iii) According to the Coverage

According to the number of units covered it can be **Macro study** or **Micro study**. Macro study is a study of the whole where as Micro study is a study of the part. For example, working capital management in State Road Transport Corporations in India is a macro study where as Working Capital Management in Andhra Pradesh State Road Transport Corporation is a micro study.

iv) According to Utility or Application

Depending upon the use of research results i.e., whether it is contributing to the theory building or problem solving, research can be Basic or Applied. **Basic research** is called pure / theoretical / fundamental research. Basic research includes original investigations for the advancement of knowledge that does not have specific objectives to answer problems of sponsoring agencies.

Applied research also called Action research, constitutes research activities on problems posed by sponsoring agencies for the purpose of contributing to the solution of these problems.

v) According to the place where it is carried out

Depending upon the place where the research is carried out (according to the data generating source), research can be classified into:

a) Field Studies or field experiments

b) Laboratory studies or Laboratory experiments

c) Library studies or documentary research

vi) According to the Research Methods used

Depending upon the research method used for the investigation, it can be classified as:

a) Survey research, b) Observation research , c) Case research, d) Experimental research, e) Historical research, f) Comparative research.

vii) **According to the Time Frame**

Depending upon the time period adopted for the study, it can be

a) **One time or single time period research -** eg. One year or a point of time. Most of the sample studies, diagnostic studies are of this type.

b) **Longitudinal research** - eg. several years or several time periods (a time series analysis) eg. industrial development during the five year plans in India.

viii) **According to the purpose of the Study**

What is the purpose/aim/objective of the study ? Is it to describe or analyze or evaluate or explore? Accordingly the studies are known as.

a) **Descriptive Study:** The major purpose of descriptive research is the description of a person, situation, institution or an event as it exists. Generally fact finding studies are of this type.

b) **Analytical Study:** The researcher uses facts or information already available and analyses them to make a critical examination of the material. These are generally Ex-post facto studies or post-mortem studies.

c) **Evaluation Study:** This type of study is generally conducted to examine / evaluate the impact of a particular event, eg. Impact of a particular decision or a project or an investment.

d) **Exploratory Study:** The information known on a particular subject matter is little. Hence, a study is conducted to know more about it so as to formulate the problem and procedures of the study. Such a study is called exploratory/ formulative study.

Self Assessment Exercise B

1) Distinguish between inductive and deductive logic.

....

....

....

2) What is the role of R & D in business?

....

....

....

3) How does research influence business decisions?

....

....

....

4) Distinguish between qualitative and quantitative data.

....

....

....

5) List the various types of studies according to the purpose of the study.

....

....

....

1.8 METHODS OF RESEARCH

The researcher has to provide answers at the end, to the research questions raised in the beginning of the study. For this purpose he has investigated and gathered the relevant data and information as a basis or evidence. The procedures adopted for obtaining the same are described in the literature as methods of research or approaches to research. In fact, they are the broad methods used to collect the data. These methods are as follows:

1) Survey Method

2) Observation Method

3) Case Method

4) Experimental Method

5) Historical Method

6) Comparative Method

It is now proposed to explain briefly, each of the above mentioned methods.

1.8.1 Survey Method

The dictionary meaning of 'Survey' is to oversee, to look over, to study, to systematically investigate. Survey research is used to study large and small populations (or universes). It is a fact finding survey. Mostly empirical problems are investigated by this approach. It is a critical inspection to gather information, often a study of an area with respect to a certain condition or its prevalence. For example: a marketing survey, a household survey, All India Rural Credit Survey.

Survey is a very popular branch of social science research. Survey research has developed as a separate research activity alongwith the development and improvement of sampling procedures. Sample surveys are very popular now a days. As a matter of fact sample survey has become synonymous with survey. For example, see the following definitions:

Survey research can be defined as **"Specification of procedures for gathering information about a large number of people by collecting information from a few of them".** (Black and Champion).

Survey research is **"Studying samples chosen from populations to discover the relative incidence, distribution, and inter relations of sociological and psychological variables".** (Fred N. Kerlinger)

By surveying data, information may be collected by observation, or personal interview, or mailed questionnaires, or administering schedules or telephone enquiries.

Features of Survey method

The important features of survey method are as follows:

i)   It is a field study, as it is always conducted in a natural setting.

ii) It solicits responses directly from the respondents or people known to have knowledge about the problem under study.

iii) Generally, it gathers information from a large population.

iv) A survey covers a definite geographical area eg. A village / city or a district.

v)   It has a time frame.

vi) It can be an extensive survey involving a wider sample or it can be an intensive study covering few samples but is an in-depth and detailed study.

vii) Survey research is best adapted for obtaining personal, socio-economic facts, beliefs, attitudes, opinions.

Survey research is not a clerical routine of gathering facts and figures. It requires a good deal of research knowledge and sophistication. The competent survey investigator must know sampling procedures, questionnaire / schedule / opionionaire construction, techniques of interviewing and other technical aspects of the survey. Ultimately the quality of the Survey results depends on the imaginative planning, representative sampling, reliability of data, appropriate analysis and interpretation of the data.

1.8.2   Observation Method

Observation means seeing or viewing. It is not a casual but systematic viewing. Observation may therefore be defined as "a systematic viewing of a specific phenomenon in its proper setting for the purpose of gathering information for the specific study".

Observation is a method of scientific enquiry. We observe a person or an event or a situation or an incident. The body of knowledge of various sciences such as biology, physiology, astronomy, sociology, psychology, anthropology etc., has been built upon centuries of systematic observation.

Observation is also useful in social and business sciences for gathering information and conceptualizing the same. For example, What is the life style of tribals? How are the marketing activities taking place in Regulated markets? How will the investment activities be done in Stock Exchange Markets? How are proceedings taking place in the Indian Parliament or Assemblies? How is a corporate office maintained in a public sector or a private sector undertaking? What is the behaviour of political leaders? Traffic jams in Delhi during peak hours?

Observation as a method of data collection has some features:

i)   It is not only seeing & viewing but also hearing and perceiving as well.

ii) It is both a physical and a mental activity. The observing eye catches many things which are sighted, but attention is also focused on data that are relevant to the problem under study.

iii) It captures the natural social context in which the person's behaviour occurs.

iv) Observation is selective: The investigator does not observe every thing but selects the range of things to be observed depending upon the nature, scope and objectives of the study.

v) Observation is not casual but with a purpose. It is made for the purpose of noting things relevant to the study.

vi) The investigator first of all observes the phenomenon and then gathers and accumulates data.

Observation may be classified in different ways. According to the setting it can be   (a) observation in a natural setting, eg. Observing the live telecast of parliament proceedings or watching from the visitors gallery, Electioneering in India through election meetings or (b) observation in an artificially stimulated setting, eg. business games, Tread Mill Test. According to the mode of observation it may be classified as (a) direct or personal observation, and (b) indirect or mechanical observation. In case of direct observation, the investigator personally observes the event when it takes place, where as in case of indirect observation it is done through mechanical devices such as audio recordings, audio visual aids, still photography, picturization etc. According to the participating role of the observer, it can be classified as (a) participant observation and (b) non-participant observation. In case of participant observation, the investigator takes part in the activity, i.e. he acts both as an observer as well as a participant. For example, studying the customs and life style of tribals by living / staying with them. In case of non-participant observation, the investigator observes from outside, merely as an on looker.

Observation method is suitable for a variety of research purposes such as a study of human behaviours, behaviour of social groups, life styles, customs and traditions, inter personal relations, group dynamics, crowd behaviour, leadership and management styles, dressing habits of different social groups in different seasons, behaviour of living creatures like birds, animals, lay out of a departmental stores, a factory or a residential locality, or conduct of an event like a meeting or a conference or Afro- Asian Games.

1.8.3   Case Method

Case method of study is borrowed from Medical Science. Just like a patient, the case is intensively studied so as to diagnose and then prescribe a remedy. A firm, or a unit is to be studied intensively with a view to finding out problems, differences, specialties so as to suggest remedial measures. It is an in-depth/intensive study of a unit or problem under study. It is a comprehensive study of a firm or an industry, or a social group, or an episode, or an incident, or a process, or a programme, or an institution or any other social unit.

According to P.V. Young "a comprehensive study of a social unit, be that unit a person, a group, a social institution, a district, or a community, is called a Case Study".

Case Study is one of the popular research methods. A case study aims at studying every thing about something rather than something about everything. It examines complex factors involved in a given situation so as to identify causal factors operating in it. The case study describes a case in terms of its peculiarities, typical or extreme features. It also helps to secure a fund of information about the unit under study. It is a most valuable method of study for diagnostic therapeutic purposes.

1.8.4   Experimental Method

Experimentation is the basic tool of the physical sciences like Physics, Chemistry for establishing cause and effect relationship and for verifying inferences. However, it is now also used in social sciences like Psychology,

Sociology. Experimentation is a research process used to observe cause and effect relationship under controlled conditions. In other words it aims at studying the effect of an independent variable on a dependent variable, by keeping the other interdependent variables constant through some type of control. In experimentation, the researcher can manipulate the independent variables and measure its effect on the dependent variable. The main features of the experimental method are :

i)   Isolation of factors or controlled observation.

ii) Replication of the experiment i.e. it can be repeated under similar conditions.

iii) Quantitative measurement of results.

iv) Determination of cause and effect relationship more precisely.

Three broad types of experiments are:

a) The natural or uncontrolled experiment as in case of astronomy made up mostly of observations.

b) The field experiment, the best suited one for social sciences. "A field experiment is a research study in a realistic situation in which one or more independent variables are manipulated by the experimenter under as carefully controlled conditions as the situation will permit". (Fred N. Kerlinger)

c) The laboratory experiment is the exclusive domain of the physical scientist.

"A laboratory experiment is a research study in which the variance of all or nearly all of the possible influential independent variables, not pertinent to the immediate problem of the investigation, is kept at a minimum. This is done by isolating the research in a physical situation apart from the routine of ordinary living and by manipulating one or more independent variables under rigorously specified, operationalized, and controlled conditions". (Fred N. Kerlinger).

The contrast between the field experiment and laboratory experiment is not sharp, the difference is a matter of degree. The laboratory experiment has a maximum of control, where as the field experiment must operate with less control.

1.8.5   Historical Method

When research is conducted on the basis of historical data, the researcher is said to have followed the historical approach. To some extent, all research is historical in nature, because to a very large extent research depends on the observations / data recorded in the past. Problems that are based on historical records, relics, documents, or chronological data can conveniently be investigated by following this method. Historical research depends on past observations or data and hence is non-repetitive, therefore it is only a post facto analysis. However, historians, philosophers, social psychiatrists, literary men, as well as social scientists use the historical approach.

Historical research is the critical investigation of events, developments, experiences of the past, the careful weighing of evidence of the validity of the sources of information of the past, and the interpretation of the weighed evidence. The historical method, also called historiography, differs from other methods in its rather elusive subject matter i.e. the past.

In historical research primary and also secondary sources of data can be used. A **primary source** is the original repository of a historical datum, like an original record kept of an important occasion, an eye witness description of an event, the inscriptions on copper plates or stones, the monuments and relics, photographs, minutes of organization meetings, documents. A **secondary source** is an account or record of a historical event or circumstance, one or more steps removed from an original repository. Instead of the minutes of the meeting of an organization, for example, if one uses a newspaper account of the meeting, it is a secondary source.

The aim of historical research is to draw explanations and generalizations from the past trends in order to understand the present and to anticipate the future. It enables us to grasp our relationship with the past and to plan more intelligently for the future.

For historical data only authentic sources should be depended upon and their authenticity should be tested by checking and cross checking the data from as many sources as possible. Many a times it is of considerable interest to use Time Series Data for assessing the progress or for evaluating the impact of policies and initiatives. This can be meaningfully done with the help of historical data.

1.8.6 Comparative Method

The comparative method is also frequently called the evolutionary or Genetic Method. The term comparative method has come about in this way: Some sciences have long been known as "Comparative Sciences" - such as comparative philology, comparative anatomy, comparative physiology, comparative psychology, comparative religion etc. Now the method of these sciences came to be described as the "Comparative Method", an abridged expression for "the method of the comparative sciences". When the method of most comparative sciences came to be directed more and more to the determination of evolutionary sequences, it came to be described as the "Evolutionary Method".

The origin and the development of human beings, their customs, their institutions, their innovations and the stages of their evolution have to be traced and established. The scientific method by which such developments are traced is known as the Genetic method and also as the Evolutionary method. The science which appears to have been the first to employ the Evolutionary method is comparative philology. It is employed to "compare" the different languages in existence, to trace the history of their evolution in the light of such similarities and differences as the comparisons disclosed. Darwin's famous work "Origin of Species" is the classic application of the Evolutionary method in comparative anatomy.

The whole theory of biological evolution rests on applications of evolutionary method. This method can be applied not only to plants, to animals, to social customs and social institutions, to the human mind (comparative psychology), to human ideas and ideals, but also to the evolution of geological strata, to the differentiation of the chemical elements and to the history of the solar system.

The term comparative method as a method of research is used here in its restricted meaning as synonymous with Evolutionary method. To say that the comparative method is a 'method of comparison' is not convincing, for comparison is not a specific method, but some thing which enters as a factor

into every scientific method. Classification requires careful comparison and every other method of science depends upon a precise comparison of phenomena and the circumstances of their occurrence. All methods are, therefore, "comparative" in a wider sense.

1.9 DIFFICULTIES IN BUSINESS RESEARCH

In India, researchers in general, and business researchers in particular are facing several problems. This is all the more true in case of empirical research. Some of the important problems are as follows:

i)   The lack of scientific training in the business research methodology is a major problem in our country. Many researchers take a leap in the dark without having a grip over research methodology. Systematic training in business research methodology is a necessity.

ii)   There is paucity of competent researchers and research supervisors. As a result the research results many a time do not reflect the reality.

iii)   Many of the business organizations are not research conscious and feel that investment in research is a wastage of resources and does not encourage research.

iv)   The research and Development Department has become a common feature in many medium and large organizations. But decision makers do not appear to be very keen on implementing the findings of their R & D departments.

v)   At the same time, small organizations which are the majority in our economy, are not able to afford a R & D department at all. Even engaging a consultant seems to be costly for them. Consequently, they do not take the help of research to solve their problems.

vi)   Many people largely depend on customs, traditions and routine practices in their decision making, as they feel that research does not have any useful purpose to serve in the management of their business.

vii)   There are insufficient interactions between the University departments and business organizations, government departments and research organizations. There should be some mechanism to develop university and industry interaction so that both can benefit i.e. the academics can get ideas from the practitioners on what needs to be researched upon and the practioners can apply the research results of the academics.

viii) The secrecy of business information is sacrosanct to business organizations. Most of the business organizations in our country do not part with information to researchers, except public sector organizations which have the culture of encouraging research, many of the private sector organizations are not willing to provide the data.

ix)   Even when research studies are undertaken, many a time, they are overlapping, resulting in duplication because there is no proper coordination between different departments of a university and between different universities.

x)   Difficulty of funds. Because of the scarcity of resources many university departments do not come forward to undertake research.

xi) Poor library facilities at many places, because of which researchers have to spend much of their time and energy in tracing out the relevant material and information.

xii) Many researchers in our country also face the difficulty of inadequate computerial and secretarial assistance, because of which the researchers have to take more time for completing their studies.

xiii) Delayed publication of data: There is difficulty of timely availability of upto date data from published sources. The data available from published sources or governmental agencies is old. At least 2 to 3 years time lag exists as a result the data proves irrelevant.

xiv) Social Research, especially managerial research, relates to human beings and their behaviour. The observations, the data collection and the conclusions etc must be valid. There is the problem of conceptualization of these aspects.

xv) Another difficulty in the research arena is that there is no code of conduct for the researchers. There is need for developing a code of conduct for researchers to educate them about ethical aspects of research, maintaining confidentiality of information etc.

In spite of all these difficulties and problems, a business enterprise cannot avoid research, especially in the fast changing world. To survive in the market an enterprise has to continuously update itself, it has to change its attitudes, approaches, products, technology, etc., through continuous research.

Self Assessment Exercise C

1) What is meant by Survey?

....

....

....

2) Distinguish between observation and experiment.

....

....

....

3) What are the comparative sciences?

....

....

....

4) What is a Case Study?

....

....

....

5) List out five important difficulties faced by business researchers in India.

....

....

....

1.10 BUSINESS RESEARCH PROCESS

In abstract terms research is research everywhere and the research process also is more or less the same, whether it is business research or agricultural research or educational research. Of course, here and there certain modifications may be required to suit the specified requirements of the area of research. The business research process also consists of a number of stages: Planning the research activity, execution of the plan and finally consolidation of the results of the research activity or reporting. The important activities involved in the research process are listed below:

i)   Selection of a research problem or researchable area.

ii)   Acquaintance with the current theory and knowledge and work done in that area.

iii) Definition and specification of the research problem more clearly.

iv)   Formulation of research hypothesis or at least research objectives.

v)   Identification of the sources of data.

vi)   Creation and construction of data collection instruments like Questionnaire, Schedules, Scales etc.

vii) Pre-testing of the instruments and their possible revision.

viii) Formal acquisition of data and information, through survey, observation, interview etc.

ix)   Processing and analysis of the data.

x)   Interpretation of the data and formal write up i.e., reporting.

These aspects are dealt with in detail in the units that follow:

Specifically, aspects (i) to (iv) are covered in unit-2, aspects (v) to (viii) are covered in units 3,4 and 5, processing and presentation aspects of (ix) are discussed in units 6 & 7, and analytical tools and techniques of data analysis of (ix) are elaborated in units 8 to 17, interpretation aspects of (x) are discussed in unit 18 and reporting aspects in unit 19. Therefore, the above aspects are not elaborated in this unit.

1.11 LET US SUM UP

Research is a part of any systematic knowledge. It is essentially a systematic investigation to discover answers to problems, seeking facts / truth. The word Science can be understood in two senses.

– Science as an organized body of knowledge and science as a method leading to knowledge. All sciences are knowledge, but all knowledge is not science.

Empirical studies have a great potential for they lead to inductions and deductions. Induction is the process of reasoning to arrive at generalizations

from particular facts. Deduction is a way of making a particular inference from a generalization.

Research is very useful to business organizations and their managers in a number of ways. It facilitates timely and objective decisions. It helps in solving business problems. It helps in providing answers to many business questions. It is of immense use to business in its functional areas. Marketing research, personnel research, production management research, financial management research, accounting research are examples.

Research can be classified into different types for the sake of better understanding. Several bases can be used for this classification such as branch of knowledge, nature of data, coverage, application, place of research, research methods used, time frame etc., and the research may be known as that type.

The research has to provide answers to the research questions raised. For this the problem has to be investigated and relevant data has to be gathered. The procedures adopted for obtaining the data and information are described as methods of research. There are six methods viz., Survey, Observation, Case, Experimental, Historical and Comparative methods.

Survey is a fact finding enquiry conducted in a natural setting/field, soliciting responses from people known to have knowledge about the problem under study. Observation is a systematic viewing of a specific phenomenon in its proper setting for gathering information. A comprehensive or in-depth study of an element of research is called a case study. Experimentation is a research process used to observe cause and effect relationship under controlled conditions. Historical research depends on past observations or past data and hence is a post facto analysis. The comparative method is an evolutionary method employed to trace the evolution, similarities and differences between the elements under study.

The business researcher in India has to face certain difficulties such as lack of scientific research training, paucity of competent researchers and research supervisors, non-encouragement of research by business organizations, small business organizations are not able to afford R & D departments, lack of scientific orientation in business management, insufficient interaction between industry and university, funding problems, poor library facilities, delayed availability of published data etc.

The business research process involves a number of stages such as selection of a researchable problem, review of previous work on that problem, specification of the problem, formulation of hypotheses / objectives, identifying sources of data, construction of data collection instruments and their pre-testing, collection of data, processing and analysis of data and finally interpretation and Report writing.

1.12 KEY WORDS

Deduction : It is a way of making a particular inference from a generalization.

Empirical : Relying/based on experience/observation/experiment

Fact : An event that is true/happened

Induction : It is a process of reasoning to arrive at generalizations from particular facts.

Knowledge : Having Information, acquaintance with facts.

Method : A way or mode of doing anything.

Observation : Systematic viewing of things to gather information.

Research : It is a systematic search for pertinent information on a specific topic.

Science : It may mean accumulated body of knowledge or it may mean a process leading to knowledge.

1.13 ANSWERS TO SELF ASSESSMENT EXERCISES

A. 1) Research is a systematic endeavour to discover answers to questions.

2) Science means Knowledge.

3) All Sciences are knowledge but all knowledge is not science.

4) A fact is a verifiable observation.

B. 1) Induction is a reasoning from particular to general, where as deduction is a reasoning from general to particular.

2) R & D helps the organization in discovery and invention.

3) By providing not only facts and figures to support decisions, but also enabling to choose one which is best.

4) Quantitative is variables based, where as qualitative is attribute based.

5) Descriptive, analytical, evaluation, exploratory studies.

C. 1) It is a fact finding from the respondents

2) Observation is an uncontrolled experiment, and experiment is a controlled observation.

3) Comparative philology, comparative anatomy, comparative religion, comparative psychology etc.

4) An intensive study of a person, a group, an incident or an institution is a case study.

1.14 TERMINAL QUESTIONS

A. Short answer Questions:

1) What do you mean by research?

2) What do you mean by Science?

3) What is knowledge?

4) What is inductive logic?

5) What is meant by deduction?

6) What are the different areas of business research?

7) What are the bases used for classifying research into different types?

8) List the various methods of research.

9) Distinguish between qualitative and quantitative data.

10) What are the stages in the business research process?

B. Essay Type Questions:

1) Define the concept of research and analyze its characteristics.

2) Define the term Science and distinguish it from knowledge.

3) Explain the significance of business research.

4) Write an essay on various types of research.

5) What do you mean by a method of research? Briefly explain different methods of research.

6) Explain the significance of research in various functional areas of business.

7) What is Survey Research? How is it different from Observation Research?

8) Write short note on:

   a) Case Research

   b) Experimental Research

   c) Historical Research

   d) Comparative Method of research

9) What are the difficulties faced by researchers of business in India?

10) What is meant by business research process? What are the various stages / aspects involved in the research process.

> **Note:** These questions/exercises will help you to understand the unit better. Try to write answers for them. But do not submit your answers to the university for assessment. These are for your practice only.

1.15   FURTHER READING

The following text books may be used for more indepth study on the topics dealt with in this unit.

Fred N. Kerlinger. *Foundations of Behavioural Research*, Surjeet Publications, Delhi

J.F.Rummel & W.C.Ballaine. *Research Methodology in Business*, Harper & Row, Publishers, Newyork

P.V.Young. *Scientific Social Surveys and Research*, Prentice-Hall of India, New Delhi

C.R.Kothari, *Research Methodology (Methods and Techniques),* New Age International Pvt. Ltd. New Delhi

T.S. Wilkinson & P.L.Bhanarkar. *Methodology and Techniques of Social Research*, Himalaya Publishing House, Mumbai

UNIT 2   RESEARCH PLAN

STRUCTURE

2.0   OBJECTIVES

After studying this unit, you should be able to:

l   Select a research problem and identify sources of research problem,

l   define and specify a research problem,

l   explain the need for formulating research objective(s),

l   define hypothesis and classify the hypotheses,

l   suggest a criteria for a good hypothesis,

l   test a hypothesis,

l   describe a research design,

l   list out the components of a research design, and

l   distinguish between a pilot study and pre-test.

2.1   INTRODUCTION

In unit 1, we have discussed the meaning and significance of business research, types of research, methods of conducting research, and the business research process. There we have shown that the research process begins with the raising of a problem, leading to the gathering of data, their analysis and interpretation and finally ends with the writing of the report. In this unit, we propose to give a complete coverage on selection and specification of the research problem, formulation of research objectives / hypotheses and designing the action plan of research. Now we will dwell in detail on these aspects along

with the associated features which are interwoven with the research problem and hypothesis formulation and testing.

2.2 RESEARCH PROBLEM

Without a problem, research cannot proceed, because there is nothing to proceed from and proceed towards. Therefore, the first step in research is to perceive a problem - either practical or theoretical. The recognition or existence of a problem motivates research. It may be noted that research is the process of repeated search for truth/facts. Unless there is a problem to search for, investigation cannot proceed. Thus, a problem sets the goal or direction of research.

A problem in simple words is "some difficulty experienced by the researcher in a theoretical or practical situation. Solving this difficulty is the task of research".

A problem exists when we do not have enough information to answer a question (problem). The answer to the question or problem is what is sought in the research.

By problem we mean "any condition or circumstance in which one does not know how to act and what to accept as true". In our common usage when we are unable to assess a thing correctly, we often say 'it is problematic'. Thus the researcher who selects a problem formulates a hypothesis or postulates a theoretical assumption that this or that is true, this or that thing to do. He/she collects proof (facts/data) of his/her hypothesis. Based on the analysis of the data collected he/she asserts the truth or answers the question/solves the problem.

The problem for research should ordinarily be expressed in an interrogative form. For example :

– Why is product X more popular than product Y?

– How to increase labour productivity?

– Does illumination increase productivity?

– Why is factory A earning profits and factory B incurring losses?

– Is the audio-visual system of teaching more effective than the audio system?

These are all searchable problems/questions. Finding answers to the problems is what is endeavoured in research. One question/problem may give rise to number of/series of sub-questions too.

Let us, now, discuss some considerations for selection of a research problems.

A topic of study may be selected by some institution or by some researcher or researchers having intellectual interests. In the former case there could be a wide variety of problems in which institutions are interested. The institution could be a local body, or government or corporate enterprises or a political party. For example, the government may be interested in assessing the probable consequences of various courses of action for solving a problem say rural unemployment. A firm may be interested in assessing the demand for something and predicting the future course of events so as to plan appropriate action relating to marketing, production, consumer behaviour and so on.

The topic of study may be selected by some individual researcher having intellectual or scientific interests. The researcher may be interested in exploring some general subject matter about which relatively little is known. And its purpose is just for scientific curiosity. Person may also be interested in a phenomenon which has already been studied in the past, but now it appears that conditions are different and, therefore, it requires further examination. Person may also be interested in a field in which there is a highly developed theoretical system but there is need for retesting the old theory on the basis of new facts, so as to test its validity in the changed circumstances.

The topic of research may be of a general nature or specifically needed by some institution, organization or government. It may be of intellectual interest or of practical concern, "A wide variety of practical concerns may present topics for research". For example, one may want to study the impact of television on children's education, performance of regulated agricultural markets, profitability of a firm, impact of imports on Indian economy, a comparative study of accounting practices in public and private undertakings, etc.

2.2.1 Sources of Research Problems

If the researcher / research organization has a ready problem on hand, he/she can proceed further in the research process or else you have to search for a problem. Where can you search for research problems? Your own mind, where else? You have to feel the problem and think about it. However, the following sources may help you in identifying the problem / problem areas.

1) **Business Problems:** A research problem is a felt need, the need may be an answer, or a solution or an improvement in facilities / technology eg. Cars Business experiences, various types of problems. They may be business policy problems, operational problems, general management problems, or functional area problems. The functional areas are Financial Management, Marketing Management, Production Management and Human Resources Management. Every business research problem is expected to solve a management problem by facilitating rational decision-making.

2) **Day to Day Problems:** A research problem can be from the day to day experience of the researcher. Every day problems constantly present some thing new and worthy of investigation and it depends on the keenness of observation and sharpness of the intellect of the researcher to knit his daily experience into a research problem. For example, a person who travels in city buses every day finds it a problem to get in or get out of the bus. But a Q system (that is the answer to the problem) facilitates boarding and alighting comfortably.

3) **Technological Changes:** Technological changes in a fast changing world are constantly bringing forth new problems and thus new opportunities for research. For example, what is the impact or implications of a new technique or new process or new machine?

4) **Unexplored Areas:** Research problems can be both abstract and of applied interest. The researcher may identify the areas in which much work has been done and the areas in which little work has been done or areas in which no work has been done. He may select those areas which have not been explored so far/explored very little.

5) **Theory of One's Own Interest:** A researcher may also select a problem for investigation from a given theory in which he has considerable interest. In such situations the researcher must have a thorough knowledge of that theory and should be able to explore some unexplained aspects or assumptions of that theory. His effort should revalidate, or modify or reject the theory.

6) **Books, Theses, Dissertation Abstracts, Articles:** Special assignments in textbooks, research theses, investigative reports, research articles in research journals etc., are rich sources for problem seekers. These sources may suggest some additional areas of needed research. Many of the research theses and articles suggest problems for further investigation which may prove fruitful.

7) **Policy Problems:** Government policy measures give rise to both positive and negative impact. The researcher may identify these aspects for his research. For example, what is the impact of the Government's new industrial policy on industrial development? What is the impact of Export - Import policy on balance of payments? What is the impact of Securities Exchange Board of India Regulations on stock markets?

8) **Discussions with Supervisor and Other Knowledgeable Persons:** The researcher may find it fruitful to have discussions with his/her proposed supervisor or other knowledgeable persons in the area of the topic.

Self Assessment Exercise A

Fill up the blanks with appropriate words

1) A research problem is a .... need.

2) The problem sets the .... of research.

3) The research problem should preferably be expressed in .... form.

4) A problem exists when we do not have enough .... to answer it.

5) Technological changes are a constant .... for research.

6) List five research problems on your own.

   ....

   ....

2.2.2 Points to be Considered while Selecting a Problem

The topic or problem which the researcher selects among the many possibilities should meet certain requirements. Every problem selected for research must satisfy the following criteria.

1) The topic selected should be original or at least less explored. The purpose of research is to fill the gaps in existing knowledge or to discover new facts and not to repeat already known facts. Therefore, a preliminary survey of the existing literature in the proposed area of research should be carried out to find out the possibility of making an original contribution. Knowledge about previous research will serve at least three purposes.

   a) It will enable the researcher to identify his specific problem for research.

   b) It will eliminate the possibility of unnecessary duplication of effort, and

   c) It will give him valuable information on the merits and limitations of various research techniques which have been used in the past.

2) It should be of significance and socially relevant and useful.

3) It should be interesting to the researcher and should fit into his aptitude.

4) It should be from an area of the researcher's specialization.

5) It should correspond to the researcher's abilities - both acquired and acquirable.

6) It should be big enough to be researchable and small enough to be handled - the topic should be amenable for research with existing and acquirable skills.

7) It should have a clear focus or objective.

8) The feasibility of carrying out research on the selected problem should be checked against the following considerations.

   a) Whether adequate and suitable data are available?

   b) Whether there is access to the organization and respondents?

   c) Whether cooperation will be forth coming from the organization and respondents?

   d) What are the resources required and how are they available?

   e) Whether the topic is within the resources (money and man power) position of the researcher?

9) It should be completed with in the time limits permissible.

2.2.3 Specification of the Problem

After going through all the above issues a problem is to be restated in an analytical jargon keeping in view its solution. The best way of understanding the problem is to discuss it with those who first raised it in order to find out how it originally came up and what was in the minds of the people who raised it. The more general the original statement of the problem, the more the necessity of preliminary discussions about its nature.

The research problem should define the goal of the researcher in clear terms. It means that along with the problem, the objective of the proposal should adequately be spelled out. Without a clear cut idea of the goal to be reached, research activities would be meaningless.

The first step in the formulation and specification of a research problem is to make it concrete and explicit. There is no foolproof method by which one can do it. However, R.L.Ackoff provides considerable guidance in identifying and specifying a problem of research. He presents five components of a problem.

1) **Research Consumer:** There must be an individual or a group which has difficulty. The individual may be the researcher himself and the group / a group of researchers. For some problems there are also other participants. The researcher, if he/she is different from the research consumer, is a participant in the problem.

2) **Research-Consumer's Objective:** The research consumer must have something to know or some ends to achieve. Obviously, a person who wants nothing cannot have a problem.

3) **Alternative Means to Achieve the Objective:** The research consumer must have alternative means to achieve his objectives. Means are courses of action. A course of action may involve the use of objects. The objects used thus are instruments. Here an instrument refers to any object, concept or idea which can be effectively used in the pursuit of an objective.

It should be remembered that there must be at least two means available to the research consumer. If he/she has no choice of means, he/she cannot have a problem.

4) **Doubt in Regard to Selection of Alternatives:** The existence of alternative courses of action is not enough. To experience a problem, the research consumer must have some doubt as to which alternative to select. Without such a doubt there can be no problem. All problems then get reduced ultimately to the evaluation of efficiency of the alternative means for a given set of objectives.

5) **One or More Environments:** There must be one or more environments to which the difficulty or problem pertains. A problem may exist in one environment and may not in another. Thus a change in environment may produce or remove a problem. A research consumer may have doubts as to which will be the most efficient means. The strategy of marketing a product may be different in the urban market, the semi-urban market and the rural market. The instruments of spreading the family planning message may be different in the case of educated and illiterate people. The range of environments over which a problem may exist may vary from one to many. Some problems are specific to only one environment while others are quite general.

The selection of a topic for research is only half-a-step forward. This general topic does not help a researcher to see what data are relevant to his/her purpose. What are the methods would he/she employ in securing them? And how to organize these? Before he/she can consider all these aspects, he/she has to formulate a specific problem by making the various components of it (as explained above) explicit.

A research problem is nothing but a basic question for which an answer or a solution is sought through research. The basic question may be further broken down into specifying questions. These "simple, pointed, limited, empirically verifiable questions are the final result of the phased process, we designate as the formulation of a research problem". Specification or definition of the problem is therefore a process that involves a progressive narrowing of the scope and sharpening of focus of questions till the specific challenging questions are finally posed. If you can answer the following questions, you have clearly specified/defined the problem.

1)   What do you want to know? (What is the problem / what are the questions to be answered).

2)   Why do you want to know? (What is the purpose or objective).

3)   How do you want to answer or solve it? (What is the methodology we want to adopt to solve it)

4)   When do you want to solve it? (Within what time limits)

5)   Where do you want to solve it? (Within what spatial limits)

6)   Who is your research-consumer? (to whom are you answering)

Please remember that a problem well put is half solved.

2.3   FORMULATION OF OBJECTIVES

Having selected and specified the research problem, the next step is to formulate the objectives of research. Research is not for the sake of research. It is undertaken to achieve some thing. Thus, research is a goal-oriented

activity. We have to identify the goal / goals to be achieved and they must be specified in order to give direction to the research study. Hence, formulation of research objectives is equally important. Once research objectives are stated, then the entire research activity will be geared to achieving those objectives. For example, we intend to examine the working of a Regulated Agricultural Market in a town to know whether it is fulfilling the objectives for which it has been set up. For this study, we will gather all the relevant information/data such as arrivals of different commodities, sources and uses of funds, facilities provided in the market, users opinions etc. Similarly, if we are clear about what we want from the research exercise, then the rest of the things will depend upon the objectives such as identifying sources of data, instruments of collection of data, tools of analyzing data. However, the objectives of the study must be clear, specific and definite.

Self Assessment Exercise B

1) List any five points which will weigh in selecting a problem.

....

....

....

2) What do you mean by specification of a problem?

....

....

....

3) What is the need for formulation of objectives of research?

....

....

....

4) How do day-to-day problems give rise to research?

....

....

....

5) What is the need for knowledge about previous research?

....

....

....

2.4   HYPOTHESIS

We know that research begins with a problem or a felt need or difficulty. The purpose of research is to find a solution to the difficulty. It is desirable that the researcher should propose a set of suggested solutions or explanations of the difficulty which the research proposes to solve. Such tentative solutions formulated as a proposition are called **hypotheses**. The suggested solutions formulated as hypotheses may or may not be the real solutions to the problem. Whether they are or not is the task of research to test and establish.

2.4.1   Meaning of Hypothesis

To understand the meaning of a hypothesis, let us see some definitions:

"A hypothesis is a tentative generalization, the validity of which remains to be tested. In its most elementary stage the hypothesis may be any guess, hunch, imaginative idea, which becomes the basis for action or investigation". (G.A.Lundberg)

"It is a proposition which can be put to test to determine validity". (Goode and Hatt).

"A hypothesis is a question put in such a way that an answer of some kind can be forth coming" - (Rummel and Ballaine).

These definitions lead us to conclude that a hypothesis is a tentative solution or explanation or a guess or assumption or a proposition or a statement to the problem facing the researcher, adopted on a cursory observation of known and available data, as a basis of investigation, whose validity is to be tested or verified.

2.4.2   Types of Hypothesis

Hypotheses can be classified in a variety of ways into different types or kinds. The following are some of the types of hypotheses:

i)   **Explanatory Hypothesis:** The purpose of this hypothesis is to explain a certain fact. All hypotheses are in a way explanatory for a hypothesis is advanced only when we try to explain the observed fact. A large number of hypotheses are advanced to explain the individual facts in life. A theft, a murder, an accident are examples.

ii) **Descriptive Hypothesis:** Some times a researcher comes across a complex phenomenon. He/ she does not understand the relations among the observed facts. But how to account for these facts? The answer is a descriptive hypothesis. A hypothesis is descriptive when it is based upon the points of resemblance of some thing. It describes the **cause** and **effect** relationship of a phenomenon e.g., the current unemployment rate of a state exceeds 25% of the work force. Similarly, the consumers of local made products constitute a significant market segment.

iii) **Analogical Hypothesis:** When we formulate a hypothesis on the basis of similarities (analogy), it is called an analogical hypothesis e.g., families with higher earnings invest more surplus income on long term investments.

iv) **Working hypothesis:** Some times certain facts cannot be explained adequately by existing hypotheses, and no new hypothesis comes up. Thus, the investigation is held up. In this situation, a researcher formulates a hypothesis which enables to continue investigation. Such a hypothesis, though inadequate and formulated for the purpose of further investigation only, is called a working hypothesis. It is simply accepted as a starting point in the process of investigation.

v)   **Null Hypothesis:** It is an important concept that is used widely in the sampling theory. It forms the basis of many tests of significance. Under this type, the hypothesis is stated negatively. It is null because it may be nullified, if the evidence of a random sample is unfavourable to the hypothesis. It is a hypothesis being tested (H_0). If the calculated value of the test is less than the permissible value, Null hypothesis is accepted, otherwise it is rejected. The rejection of a null hypothesis implies that the difference could not have arisen due to chance or sampling fluctuations.

vi) **Statistical Hypothesis:** Statistical hypotheses are the statements derived from a sample. These are quantitative in nature and are numerically measurable. For example, the market share of product X is 70%, the average life of a tube light is 2000 hours etc.

2.4.3 Criteria for Workable Hypothesis

A hypothesis controls and directs the research study. When a problem is felt, we require the hypothesis to explain it. Generally, there is more than one hypothesis which aims at explaining the same fact. But all of them cannot be equally good. Therefore, how can we judge a hypothesis to be true or false, good or bad? Agreement with facts is the sole and sufficient test of a true hypothesis. Therefore, certain conditions can be laid down for distinguishing a good hypothesis from bad ones. The formal conditions laid down by thinkers provide the criteria for judging a hypothesis as good or valid. These conditions are as follows:

i) **A hypothesis should be empirically verifiable:** The most important condition for a valid hypothesis is that it should be empirically verifiable. A hypothesis is said to be verifiable, if it can be shown to be either true or false by comparing with the facts of experience directly or indirectly. A hypothesis is true if it conforms to facts and it is false if it does not. Empirical verification is the characteristic of the scientific method.

ii) **A hypothesis should be relevant:** The purpose of formulating a hypothesis is always to explain some facts. It must provide an answer to the problem which initiated the enquiry. A hypothesis is called relevant if it can explain the facts of enquiry.

iii) **A hypothesis must have predictive and explanatory power:** Explanatory power means that a good hypothesis, over and above the facts it proposes to explain, must also explain some other facts which are beyond its original scope. We must be able to deduce a wide range of observable facts which can be deduced from a hypothesis. The wider the range, the greater is its explanatory power.

iv) **A hypothesis must furnish a base for deductive inference on consequences:** In the process of investigation, we always pass from the known to the unknown. It is impossible to infer any thing from the absolutely unknown. We can only infer what would happen under supposed conditions by applying the knowledge of nature we possess. Hence, our hypothesis must be in accordance with our previous knowledge.

v) **A hypothesis does not go against the traditionally established knowledge:** As far as possible, a new hypothesis should not go against any previously established law or knowledge. The new hypothesis is expected to be consistent with the established knowledge.

vi) **A hypothesis should be simple:** A simple hypothesis is preferable to a complex one. It some times happens that there are two or more hypotheses which explain a given fact equally well. Both of them are verified by observable facts. Both of them have a predictive power and both are consistent with established knowledge. All the important conditions of hypothesis are thus satisfied by them. In such cases the simpler one is to be accepted in preference to the complex one.

vii) **A hypothesis must be clear, definite and certain:** It is desirable that the hypothesis must be simple and specific to the point. It must be clearly defined in a manner commonly accepted. It should not be vague or ambiguous.

(viii) **A Hypothesis should be related to available techniques:** If tools and techniques are not available we cannot test the hypothesis. Therefore, the hypothesis should be formulated only after due thought is given to the methods and techniques that can be used to measure the concepts and variables related to the hypothesis.

2.4.4 Stages in Hypothesis

There are four stages. The first stage is feeling of a problem. The observation and analysis of the researcher reveals certain facts. These facts pose a problem. The second stage is formulation of a hypothesis or hypotheses. A tentative supposition/ guess is made to explain the facts which call for an explanation. At this stage some past experience is necessary to pick up the significant aspects of the observed facts. Without previous knowledge, the investigation becomes difficult, if not impossible. The third stage is deductive development of hypothesis using deductive reasoning. The researcher uses the hypothesis as a premise and draws a conclusion from it. And the last stage is the verification or testing of hypothesis. This consists in finding whether the conclusion drawn at the third stage is really true. Verification consists in finding whether the hypothesis agrees with the facts. If the hypothesis stands the test of verification, it is accepted as an explanation of the problem. But if the hypothesis does not stand the test of verification, the researcher has to search for further solutions.

To explain the above stages let us consider a simple example. Suppose, you have started from your home for college on your scooter. A little while later the engine of your scooter suddenly stops. What can be the reason? Why has it stopped? From your past experience, you start guessing that such problems generally arise due to either petrol or spark plug. Then start deducing that the cause could be: (i) that the petrol knob is not on. (ii) that there is no petrol in the tank. (iii) that the spark plug has to be cleaned. Then start verifying them one after another to solve the problem. First see whether the petrol knob is on. If it is not, switch it on and start the scooter. If it is already on, then see whether there is petrol or not by opening the lid of the petrol tank. If the tank is empty, go to the near by petrol bunk to fill the tank with petrol. If there is petrol in the tank, this is not the reason, then you verify the spark plug. You clean the plug and fit it. The scooter starts. That means the problem is with the spark plug. You have identified it. So you got the answer. That means your problem is solved.

2.4.5 Testing of Hypothesis

When the hypothesis has been framed in the research study, it must be verified as true or false. Verifiability is one of the important conditions of a good hypothesis. Verification of hypothesis means testing of the truth of the hypothesis in the light of facts. If the hypothesis agrees with the facts, it is said to be true and may be accepted as the explanation of the facts. But if it does not agree it is said to be false. Such a false hypothesis is either totally rejected or modified. Verification is of two types viz., **Direct verification** and **Indirect verification**.

Direct verification may be either by observation or by experiments. When direct observation shows that the supposed cause exists where it was thought to exist, we have a direct verification. When a hypothesis is verified by an experiment in a laboratory it is called direct verification by experiment. When the hypothesis is not amenable for direct verification, we have to depend on

indirect verification. **Indirect verification** is a process in which certain possible consequences are deduced from the hypothesis and they are then verified directly. Two steps are involved in indirect verification. (i) Deductive development of hypothesis: By deductive development certain consequences are predicted and (ii) finding whether the predicted consequences follow. If the predicted consequences come true, the hypothesis is said to be indirectly verified. Verification may be done directly or indirectly or through logical methods.

Testing of a hypothesis is done by using statistical methods. Testing is used to accept or reject an assumption or hypothesis about a random variable using a sample from the distribution. The assumption is the null hypothesis (H_0), and it is tested against some alternative hypothesis (H_1). Statistical tests of hypothesis are applied to sample data. The procedure involved in testing a hypothesis is A) select a sample and collect the data. B) convert the variables or attributes into statistical form such as mean, proportion. C) formulate hypotheses. D) select an appropriate test for the data such as t-test, Z-test. E) perform computations. F) finally draw the inference of accepting or rejecting the null hypothesis. You will learn more about it in tests of hypothesis or tests of significance in later units (Units-15, 16 and 17).

2.4.6 Uses of Hypothesis

If a clear scientific hypothesis has been formulated, half of the research work is already done. The advantages/utility of having a hypothesis are summarized here underneath:

i) It is a starting point for many a research work.

ii) It helps in deciding the direction in which to proceed.

iii) It helps in selecting and collecting pertinent facts.

iv) It is an aid to explanation.

v) It helps in drawing specific conclusions.

vi) It helps in testing theories.

vii) It works as a basis for future knowledge.

Self Assessment Exercise C

1) What do you mean by a hypothesis?

....
....
....

2) List out different types of hypothesis.

....
....
....

3) What is meant by null hypothesis?

....
....
....

4) What are the characteristics of good hypothesis testing?

    ....

    ....

    ....

5) What are the stages in a hypothesis?

    ....

    ....

    ....

6) What are the methods used to prove or reject a hypothesis ?

    ....

    ....

    ....

2.5   RESEARCH DESIGN

Research design is also known by different names such as research outline, plan, blue print. In the words of Fred N. Kerlinger, it is the plan, structure and strategy of investigation conceived so as to obtain answers to research questions and control variance. The plan includes everything the investigator will do from writing the hypothesis and their operational implications to the final analysis of data. The structure is the outline, the scheme, the paradigms of the operation of the variables. The strategy includes the methods to be used to collect and analyze the data. At the beginning this plan (design) is generally vague and tentative. It undergoes many modifications and changes as the study progresses and insights into it deepen. The working out of the plan consists of making a series of decisions with respect to what, why, where, when, who and how of the research.

According to Pauline V.Young "a research design is the logical and systematic planning and directing of a piece of research". According to Reger E.Kirk "research designs are plans that specify how data should be collected and analyzed".

The research has to be geared to the available time, energy, money and to the availability of data. There is no such thing as a single or correct design. Research design represents a compromise dictated by many practical considerations that go into research.

2.5.1   Functions of Research Design

Regardless of the type of research design selected by the investigator, all plans perform one or more functions outlined below.

i)   It provides the researcher with a blue print for studying research questions.

ii) It dictates boundaries of research activity and enables the investigator to channel his energies in a specific direction.

iii) It enables the investigator to anticipate potential problems in the implementation of the study.

iv) The common function of designs is to assist the investigator in providing answers to various kinds of research questions.

A study design includes a number of component parts which are interdependent and which demand a series of decisions regarding the definitions, methods, techniques, procedures, time, cost and administration aspects.

2.5.2 Components of a Research Design

A research design basically is a plan of action. Once the research problem is selected, then it must be executed to get the results. Then how to go about it? What is its scope? What are the sources of data? What is the method of enquiry? What is the time frame? How to record the data? How to analyze the data? What are the tools and techniques of analysis? What is the manpower and organization required? What are the resources required? These and many such are the subject matter of attacking the research problem demanding decisions in the beginning itself to have greater clarity about the research study. It is similar to having a building plan before the building is constructed. Thus, according to P.V. Young the various "considerations which enter into making decisions regarding what, where, when, how much, by what means constitute a plan of study or a study design". Usually the contents or components of a Research design are as follows:

1) **Need for the Study:** Explain the need for and importance of this study and its relevance.

2) **Review of Previous Studies:** Review the previous works done on this topic, understand what they did, identify gaps and make a case for this study and justify it.

3) **Statement of Problem:** State the research problem in clear terms and give a title to the study.

4) **Objectives of Study:** What is the purpose of this study? What are the objectives you want to achieve by this study? The statement of objectives should not be vague. They must be specific and focussed.

5) **Formulation of Hypothesis:** Conceive possible outcome or answers to the research questions and formulate into hypothesis tests so that they can be tested.

6) **Operational Definitions:** If the study is using uncommon concepts or unfamiliar tools or using even the familiar tools and concepts in a specific sense, they must be specified and defined.

7) **Scope of the Study:** It is important to define the scope of the study, because the scope decides what is within its purview and what is outside.

Scope includes
- Geographical area to be covered.
- Subject content to be covered.
- Time period to be covered.

Scope includes Geographical scope, content scope, chronological scope of the study. The territorial area to be covered by the study should be decided. E.g. only Delhi or northern states or All India. As far as content scope is concerned according to the problem say for example, industrial relations in so and so organization, what are aspects to be studied, what are the aspects not coming under this and hence not studied. Chronological scope i.e., time period selection and its justification is important. Whether the study is at a point of time or longitudinal say 1991-2003.

8) **Sources of Data:** This is an important stage in the research design. At this stage, keeping in view the nature of research, the researcher has to decide the sources of data from which the data are to be collected. Basically the sources

are divided into primary source (field sources) and secondary source (documentary sources). The data from primary source are called as primary data, and data from secondary source are called secondary data. Hence, the researcher has to decided whether to collect from primary source or secondary source or both sources. (This will be discussed in detail in Unit-3).

9) **Method of Collection:** After deciding the sources for data collection, the researcher has to determine the methods to be employed for data collection, primarily, either census method or sampling method. This decision may depend on the nature, purpose, scope of the research and also time factor and financial resources.

10) **Tools & Techniques:** The tools and techniques to be used for collecting data such as observation, interview, survey, schedule, questionnaire, etc., have to be decided and prepared.

11) **Sampling Design:** If it is a sample study, the sampling techniques, the size of sample, the way samples are to be drawn etc., are to be decided.

12) **Data Analysis:** How are you going to process and analyze the data and information collected? What simple or advanced statistical techniques are going to be used for analysis and testing of hypothesis, so that necessary care can be taken at the collection stage.

13) **Presentation of the Results of Study:** How are you going to present the results of the study? How many chapters? What is the chapter scheme? The chapters, their purpose, their titles have to be outlined. It is known as chapterisation.

14) **Time Estimates:** What is the time available for this study? Is it limited or unlimited time? Generally, it is a time bound study. The available or permitted time must be apportioned between different activities and the activities to be carried out within the specified time. For example, preparation of research design one month, preparation of questionnaire one month, data collection two months, analysis of data two months, drafting of the report two months etc.,

15) **Financial Budget:** The design should also take into consideration the various costs involved and the sources available to meet them. The expenditures like salaries (if any), printing and stationery, postage and telephone, computer and secretarial assistance etc.

16) **Administration of the Enquiry:** How is the whole thing to be executed? Who does what and when? All these activities have to be organized systematically, research personnel have to be identified and trained. They must be entrusted with the tasks, the various activities are to be coordinated and the whole project must be completed as per schedule.

Research designs provide guidelines for investigative activity and not necessarily hard and fast rules that must remain unbroken. As the study progresses, new aspects, new conditions and new connecting links come to light and it is necessary to change the plan / design as circumstances demand. A universal characteristic of any research plan is its flexibility.

Depending upon the method of research, the designs are also known as survey design, case study design, observation design and experimental design.

2.6 PILOT STUDY AND PRE-TESTING

A **Pilot study** is a small scale replica of the main study. When a problem is selected for research, a plan of action is to be designed to proceed further. But if we do not have adequate knowledge about the subject matter, the nature of the population (The word 'population' as used in statistics denotes the aggregate from which the sample is to be taken), the various issues involved, the tools and techniques to be used for operationalizing the research problem, we have to familiarize ourselves first with it and acquire a good deal of knowledge about the subject matter of the study and its dimensions. For this purpose, a small study is conducted before the main study, which is called a Pilot Study. A pilot study provides a better knowledge of the problem and its dimensions. It facilitates us to understand the nature of the population to be surveyed and the field problems to be encountered. It also helps in developing better approaches and better instruments. It covers the entire process of research, but on a small scale. This is also useful for preparing the research design clearly and specifically.

Pre-Testing is the hallmark of scientific research. Pre-testing means trial administration of the instrument to sample respondents before finalizing it. It is a common practice in our day-to-day life that before finally okaying it we try it on a trial basis. For example, when some recipe is prepared a sample is tasted, based on that corrections are made. If you give a suit for stitching to the tailor, you want a trail wear (or pre-test), if you are purchasing a vehicle, you want to have a trial drive. Similarly, for data collection some instruments such as interview schedule, or questionnaire or measurement scale are constructed. We want to administer it on a trial basis to identify its weaknesses, if any. Such a trial administration of the instrument is called pre-testing.

While designing the instrument or method, we take all precautions keeping in view the requirements of the study. We will not be able to identify its defects, limitations and weaknesses easily. But when others use it, they will be able to identify them objectively. Therefore, it has to be tested empirically, hence pre-testing of a draft instrument is a must. Based on the opinions, comments, criticism, suggestions received and difficulties experienced in the pre-testing the instrument or method is revised or modified and then finalized for using it in the main study.

The difference between pilot study and pre-test is that, the former is a full fledged miniature study of a research problem, where as the latter is a trial test of a specific aspect of the study, such as a questionnaire.

Self Assessment Exercise D

1) What are the different names of research design?

   ....

   ....

   ....

2) What is meant by research plan?

   ....

   ....

   ....

3) What are the functions of a research design?

....

....

....

4) What do you mean by scope of study?

....

....

....

5) Distinguish between pilot study and pre-test.

....

....

....

2.7 LET US SUM UP

Without a problem, research cannot proceed. A problem is some difficulty experienced by the researcher in a theoretical or practical situation. Solving this difficulty is the task of research. The problem for research should ordinarily be expressed in an interrogative form. If the researcher has a ready problem on hand he can proceed further. Otherwise, he has to search for a problem. The problem can be from business in general or functional areas in particular. Other sources of research problems are: day to day problems, technological changes, unexplored areas, books, theses, articles, policy problems etc., Having selected the problem it must be defined and specified.

Having specified the problem, the next step is to formulate the objectives of research so as to give direction to the study. The researcher should also propose a set of suggested solutions to the problem under study. Such tentative solutions formulated are called hypotheses. The hypotheses are of various types such as explanatory hypothesis, descriptive hypothesis, analogical hypothesis, working hypothesis, null hypothesis and statistical hypothesis. A good hypothesis must be empirically verifiable, should be relevant, must have explanatory power, must be as far as possible within the established knowledge, must be simple, clear and definite. There are four stages in a hypothesis (a) feeling a problem (b) formulating hypothesis (c) deductive development of hypothesis and (d) verification / testing of hypothesis verification can be done either directly or indirectly or through logical methods. Testing is done by using statistical methods.

Having selected the problem, formulated the objectives and hypothesis, the researcher has to prepare a blue print or plan of action, usually called as research design. The design/study plan includes a number of components which are interdependent and which demand a series of decisions regarding definitions, scope, methods, techniques, procedures, instruments, time, place, expenditure and administration aspects.

If the problem selected for research is not a familiar one, a pilot study may be conducted to acquire knowledge about the subject matter, and the various issues involved. Then for collection of data instruments and/or scales have to constructed, which have to be pre-tested before finally accepting them for use.

2.8 KEY WORDS

Hypothesis : A hypothesis is a tentative answer / solution to the research problem, whose validity remains to be tested.

Pilot Study : A study conducted to familiarize oneself first with the research problem so that it can be operationalised with a good deal of knowledge about the problem.

Pre-Test : A trial administration of an instrument such as a questionnaire or scale to identify its weaknesses is called a pre-test.

Research Design : It is a systematic plan (planning) to direct a piece of research work.

Research Problem : A research problem is a felt need, which needs an answer/solution.

Testing of Hypothesis : It means verification of a hypothesis as true or false in the light of facts.

2.9 ANSWERS TO SELF ASSESSMENT EXERCISES

A. 1) felt   2) goal / direction

    3) interrogative   4) information / data

    5) source of problems

2.10 TERMINAL QUESTIONS

A) Short answer Questions:

1) What is meant by a research problem?

2) What do you mean by specification of the problem?

3) What is the need for formulating research objectives?

4) What do you mean by a hypothesis?

5) What are the stages in a hypothesis?

6) What do you mean by testing of a hypothesis?

7) What is a research design?

8) What are the functions of a research design?

9) What is a Pilot Study?

10) What do you mean by Pre-testing?

B) Essay Type Questions:

1) What is a research problem? Explain the sources of research problems.

2) What do you mean by a problem? Explain the various points to be considered while selecting a problem.

3) Explain how you will select and specify a research problem.

4) What do you mean by a hypothesis? What are the different types of hypotheses?

5) What is meant by hypothesis? Explain the criteria for a workable hypothesis.

6) What are the different stages in a hypothesis? How do you verify / test a hypothesis?

7) What is a research design? Explain the functions of a research design.

8) Define a research design and explain its contents.

9) What are the various components of a research design?

10) Distinguish between pilot study and pre-test. Also explain the need for Pilot study and pre-testing.

> **Note:** These questions/exercises will help you to understand the unit better. Try to write answers for them. But do not submit your answers to the university for assessment. These are for your practice only.

2.11   FURTHER READING

The following text books may be used for more indepth study on the topics dealt with in this unit.

Fred N. Kerlinger. *Foundations of Behavioural Research*, Surjeet Publications, Delhi.

O.R.Krishna Swamy. Methodology of Research in Social Sciences, Himalaya Publishing House, Mumbai.

T.S.Wilkinson & P.L.Bhandarkar. *Methodology and Techniques of Social Research*, Himalaya Publishing House, Mumbai.

C.R.Kothari. *Research Methodology*, Wiley Eastern, New Delhi.

V.P.Michael. *Research Methodology in Management*, Himalaya Publishing House, Mumbai.

UNIT 3   COLLECTION OF DATA

STRUCTURE

3.0   OBJECTIVES

On the completion of this unit, you should be able to:

l   discuss the necessity and usefulness of data collection,

l   explain and distinguish between primary data and secondary data,

l   explain the sources of secondary data and its merits and demerits,

l   describe different methods of collecting primary data and their merits and demerits,

l   examine the choice of a suitable method, and

l   examine the reliability, suitability and adequacy of secondary data.

3.1   INTRODUCTION

In Unit 2, we have discussed about the selection of a research problem and formulation of research design. A research design is a blue print which directs the plan of action to complete the research work. As we have mentioned earlier, the collection of data is an important part in the process of research work. The quality and credibility of the results derived from the application of research methodology depends upon the relevant, accurate and adequate data. In this unit, we shall study about the various sources of data and methods of collecting primary and secondary data with their merits and limitations and also the choice of suitable method for data collection.

3.2   MEANING AND NEED FOR DATA

Data is required to make a decision in any business situation. The researcher is faced with one of the most difficult problems of obtaining suitable, accurate and adequate data. Utmost care must be exercised while collecting data because

the quality of the research results depends upon the reliability of the data. Suppose, you are the Director of your company. Your Board of Directors has asked you to find out why the profit of the company has decreased since the last two years. Your Board wants you to present facts and figures. What are you going to do?

The first and foremost task is to collect the relevant information to make an analysis for the above mentioned problem. It is, therefore, the information collected from various sources, which can be expressed in quantitative form, for a specific purpose, which is called data. The rational decision maker seeks to evaluate information in order to select the course of action that maximizes objectives. For decision making, the input data must be appropriate. This depends on the appropriateness of the method chosen for data collection. The application of a statistical technique is possible when the questions are answerable in quantitative nature, for instance; the cost of production, and profit of the company measured in rupees, age of the workers in the company measured in years. Therefore, the first step in statistical activities is to gather data. The data may be classified as primary and secondary data. Let us now discuss these two kinds of data in detail.

3.3 PRIMARY AND SECONDARY DATA

The **Primary data** are original data which are collected for the first time for a specific purpose. Such data are published by authorities who themselves are responsible for their collection. The **Secondary data** on the other hand, are those which have already been collected by some other agency and which have already been processed. Secondary data may be available in the form of **published** or **unpublished** sources. For instance, population census data collected by the Government in a country is primary data for that Government. But the same data becomes secondary for those researchers who use it later. In case you have decided to collect primary data for your investigation, you have to identify the sources from where you can collect that data. For example, if you wish to study the problems of the workers of X Company Ltd., then the workers who are working in that company are the source. On the other hand, if you have decided to use secondary data, you have to identify the secondary source who have already collected the related data for their study purpose.

With the above discussion, we can understand that the difference between primary and secondary data is only in terms of degree. That is that the data which is primary in the hands of one becomes secondary in the hands of another.

Self Assessment Exercise A

1) What do you mean by data? Why it is needed for research?

    ....
    ....
    ....

2) Distinguish between primary and secondary data. Illustrate your answer with examples.

    ....
    ....
    ....

3.4 SOURCES OF SECONDARY DATA

We have discussed above the meaning of primary and secondary data. Sometimes, it is not possible to collect primary data due to time, cost and human resource constraints. Therefore, researchers have to take the help of secondary data. Now let us discuss, (a) various sources from where, one can get secondary data, (b) precautions while using secondary data, its merits and demerits and some documentary and electronic sources of data in India.

3.4.1 Documentary Sources of Data

This category of secondary data source may also be termed as **Paper Source**. The main sources of documentary data can be broadly classified into two categories:

a) Published sources, and

b) Unpublished sources.

Let us discuss these two categories in detail.

a) **Published Sources**

There are various national and international institutions, semi-official reports of various committees and commissions and private publications which collect and publish statistical data relating to industry, trade, commerce, health etc. These publications of various organisations are useful sources of secondary data. These are as follows:

1) **Government Publications:** Central and State Governments publish current information alongwith statistical data on various subjects, quarterly and annually. For example, Monthly Statistical Abstract, National Income Statistics, Economic Survey, Reports of National Council of Applied Economic Research (NCEAR), Federation of Indian Chambers of Commerce and Industry (FICCI), Indian Council of Agricultural Research (ICAR), Central Statistical Organisation (CSO), etc.

2) **International Publications:** The United Nations Organisation (UNO), International Labour Organisation (ILO), International Monetary Fund (IMF), World Bank, Asian Development Bank (ADB) etc., also publish relevant data and reports.

3) **Semi-official Publications:** Semi-official organisations like Corporations, District Boards, Panchayat etc. publish reports.

4) **Committees and Commissions:** Several committees and commissions appointed by State and Central Governments provide useful secondary data. For example, the report of the 10th Financial Commission or Fifth Pay Commissions etc.

5) **Private Publications:** Newspapers and journals publish the data on different fields of Economics, Commerce and Trade. For example, Economic Times, Financial Express etc. and Journals like Economist, Economic and Political Weekly, Indian Journal of Commerce, Journal of Industry and Trade, Business Today etc. Some of the research and financial institutions also publish their reports annually like Indian Institute of Finance. In addition to this, reports prepared by research scholars, universities etc. also provide secondary source of information.

b) Unpublished Sources

It is not necessary that all the information/data maintained by the institutions or individuals are available in published form. Certain research institutions, trade associations, universities, research scholars, private firms, business institutions etc., do collect data but they normally do not publish it. We can get this information from their registers, files etc.

3.4.2 Electronic Sources

The secondary data is also available through electronic media (through Internet). You can download data from such sources by entering web sites like *google.com; yahoo.com; msn.com;* etc., and typing your subject for which the information is needed.

You can also find secondary data on electronic sources like CDs, and the following online journals:

Electronic Journal	http://businessstandard.com
Electronic Journal	http://www.businessworldindia.com
Electronic Journal	http://www.business-today.com
Electronic Journal	http://www.india-invest.com
Census of India	http://www.censusindia.net
Union Budget and Economic Survey	http://www.indianbudget.nic.in
Directory of Government of India Institutions	http://goidirectory.nic.in
Indian Council of Agricultural Research	http://www.icar.org.in
Ministry of Commerce and Industry	http://www.commin.nic.in
Indian Institute of Foreign Trade	http://www.iift.edu
Department of Industrial Policy and Promotion, Ministry of Commerce and Industry	http://www.dipp.nic.in
Ministry of Consumer Affairs, Food & Public Distribution	http://www.fccimin.in
Khadi and Village Industries	http://www.kvic.org.in
Board for Industrial & Financial Reconstruction	http://www.bifr.nic.in
Building Material & Technology Promotion Council	http://www.bmtpc.org
Central Food Technological Research Institute	http://www.cftri.com
National Council for Traders Information	http://www.ncti-india.com
National Handloom Development Corporation Ltd.	http://www.nhdcltd.com
The Associated Chamber of Commerce and Industry	http://www.assochm.org
Federation of Indian Chambers of Commerce and Industry	http://www.ficiofindia.com

Now you have learnt that the secondary data are available in documents, either published or unpublished, and electronic sources. However, you have to take precautions while using secondary data in research. Let us discuss them in detail.

3.4.3 Precaution in Using Secondary Data

With the above discussion, we can understand that there is a lot of published and unpublished sources where researcher can gets secondary data. However, the researcher must be cautious in using this type of data. The reason is that such type of data may be full of errors because of bias, inadequate size of the sample, errors of definitions etc. Bowley expressed that it is never safe to take published or unpublished statistics at their face value without knowing their meaning and limitations. Hence, before using secondary data, you must examine the following points.

Suitability of Secondary Data

Before using secondary data, you must ensure that the data are suitable for the purpose of your enquiry. For this, you should compare the objectives, nature and scope of the given enquiry with the original investigation. For example, if the objective of our enquiry is to study the salary pattern of a firm including perks and allowances of employees. But, secondary data is available only on basic pay. Such type of data is not suitable for the purpose of the study.

Reliability of Secondary Data

For the reliability of secondary data, these can be tested: i) unbiasedness of the collecting person, ii) proper check on the accuracy of field work, iii) the editing, tabulating and analysis done carefully, iv) the reliability of the source of information, v) the methods used for the collection and analysis of the data. If the data collecting organisations are government, semi-government and international, the secondary data are more reliable corresponding to data collected by individual and private organisations.

Adequacy of Secondary Data

Adequacy of secondary data is to be judged in the light of the objectives of the research. For example, our objective is to study the growth of industrial production in India. But the published report provide information on only few states, then the data would not serve the purpose. Adequacy of the data may also be considered in the light of duration of time for which the data is available. For example, for studying the trends of per capita income of a country, we need data for the last 10 years, but the information available for the last 5 years only, which would not serve our objective.

Hence, we should use secondary data if it is reliable, suitable and adequate.

3.4.4 Merits and Limitations of Secondary Data

Merits

1) Secondary data is much more economical and quicker to collect than primary data, as we need not spend time and money on designing and printing data collection forms (questionnaire/schedule), appointing enumerators, editing and tabulating data etc.

2) It is impossible to an individual or small institutions to collect primary data with regard to some subjects such as population census, imports and exports of different countries, national income data etc. but can obtain from secondary data.

Limitations

1) Secondary data is very risky because it may not be suitable, reliable, adequate and also difficult to find which exactly fit the need of the present investigation.

2) It is difficult to judge whether the secondary data is sufficiently accurate or not for our investigation.

3) Secondary data may not be available for some investigations. For example, bargaining strategies in live products marketing, impact of T.V. advertisements on viewers, opinion polls on a specific subject, etc. In such situations we have to collect primary data.

Self Assessment Exercise B

1) Write names of five web sources of secondary data which have not been included in the above table.

....

....

....

2) Explain the merits and limitations of using secondary data.

....

....

....

3) What precautions must a researcher take before using the secondary data?

....

....

....

4) In the following situations indicate whether data from a census should be taken?

   i)   A TV manufacturer wants to obtain data on customer preferences with respect to size of TV.

   ii) IGNOU wants to determine the acceptability of its employees for subscribing to a new employee insurance programme.

   ....

   ....

   ....

3.5   METHODS OF COLLECTING PRIMARY DATA

If the available secondary data does not meet the requirements of the present study, the researcher has to collect primary data. As mentioned earlier, the data which is collected for the first time by the researcher for his own purpose is

called primary data. There are several methods of collecting primary data, such as observation, interview through reporters, questionnaires and schedules. Let us study about them in detail.

3.5.1   Observation Method

The Concise Oxford Dictionary defines observation as, 'accurate watching and noting of phenomena as they occur in nature with regard to cause and effect or mutual relations'. Thus observation is not only a systematic watching but it also involves listening and reading, coupled with consideration of the seen phenomena. It involves three processes. They are: sensation, attention or concentration and perception.

Under this method, the researcher collects information directly through observation rather than through the reports of others. It is a process of recording relevant information without asking anyone specific questions and in some cases, even without the knowledge of the respondents. This method of collection is highly effective in behavioural surveys. For instance, a study on behaviour of visitors in trade fairs, observing the attitude of workers on the job, bargaining strategies of customers etc. Observation can be participant observation or non-participant observation. In **Participant Observation Method**, the researcher joins in the daily life of informants or organisations, and observes how they behave. In the **Non-participant Observation Method**, the researcher will not join the informants or organisations but will watch from outside.

Merits

1) This is the most suitable method when the informants are unable or reluctant to provide information.

2) This method provides deeper insights into the problem and generally the data is accurate and quicker to process. Therefore, this is useful for intensive study rather than extensive study.

Limitations

Despite of the above merits, this method suffers from the following limitations:

1) In many situations, the researcher cannot predict when the events will occur. So when an event occurs there may not be a ready observer to observe the event.

2) Participants may be aware of the observer and as a result may alter their behaviour.

3) Observer, because of personal biases and lack of training, may not record specifically what he/she observes.

4) This method cannot be used extensively if the inquiry is large and spread over a wide area.

3.5.2   Interview Method

Interview is one of the most powerful tools and most widely used method for primary data collection in business research. In our daily routine we see interviews on T.V. channels on various topics related to social, business, sports, budget etc. In the words of C. William Emory, 'personal interviewing is a two-way purposeful conversation initiated by an interviewer to obtain information that is relevant to some research purpose'. Thus an interview is basically, a

meeting between two persons to obtain the information related to the proposed study. The person who is interviewing is named as interviewer and the person who is being interviewed is named as informant. It is to be noted that, the research data/information collect through this method is not a simple conversation between the investigator and the informant, but also the glances, gestures, facial expressions, level of speech etc., are all part of the process. Through this method, the researcher can collect varied types of data intensively and extensively.

Interviewes can be classified as direct personal interviews and indirect personal interviews. Under the techniques of **direct personal interview**, the investigator meets the informants (who come under the study) personally, asks them questions pertaining to enquiry and collects the desired information. Thus if a researcher intends to collect the data on spending habits of Delhi University (DU) students, he/ she would go to the DU, contact the students, interview them and collect the required information.

Indirect personal interview is another technique of interview method where it is not possible to collect data directly from the informants who come under the study. Under this method, the investigator contacts third parties or witnesses, who are closely associated with the persons/situations under study and are capable of providing necessary information. For example, an investigation regarding a bribery pattern in an office. In such a case it is inevitable to get the desired information indirectly from other people who may be knowing them. Similarly, clues about the crimes are gathered by the CBI. Utmost care must be exercised that these persons who are being questioned are fully aware of the facts of the problem under study, and are not motivated to give a twist to the facts.

Another technique for data collection through this method can be structured and unstructured interviewing. In the **Structured interview** set questions are asked and the responses are recorded in a standardised form. This is useful in large scale interviews where a number of investigators are assigned the job of interviewing. The researcher can minimise the bias of the interviewer. This technique is also named as formal interview. In **Un-structured interview**, the investigator may not have a set of questions but have only a number of key points around which to build the interview. Normally, such type of interviews are conducted in the case of an explorative survey where the researcher is not completely sure about the type of data he/ she collects. It is also named as informal interview. Generally, this method is used as a supplementary method of data collection in conducting research in business areas.

Now-a-days, telephone or cellphone interviews are widely used to obtain the desired information for small surveys. For instance, interviewing credit card holders by banks about the level of services they are receiving. This technique is used in industrial surveys specially in developed regions.

Merits

The major merits of this method are as follows:

1) People are more willing to supply information if approached directly. Therefore, personal interviews tend to yield high response rates.

2) This method enables the interviewer to clarify any doubt that the interviewee might have while asking him/her questions. Therefore, interviews are helpful in getting reliable and valid responses.

3) The informant's reactions to questions can be properly studied.

4) The researcher can use the language of communication according to the standard of the information, so as to obtain personal information of informants which are helpful in interpreting the results.

Limitations

The limitations of this method are as follows:

1) The chance of the subjective factors or the views of the investigator may come in either consciously or unconsciously.

2) The interviewers must be properly trained, otherwise the entire work may be spoiled.

3) It is a relatively expensive and time-consuming method of data collection especially when the number of persons to be interviewed is large and they are spread over a wide area.

4) It cannot be used when the field of enquiry is large (large sample).

Precautions : While using this method, the following precautions should be taken:

ı obtain thorough details of the theoretical aspects of the research problem.

ı Identify who is to be interviewed.

ı The questions should be simple, clear and limited in number.

ı The investigator should be sincere, efficient and polite while collecting data.

ı The investigator should be of the same area (field of study, district, state etc.).

Self Assessment Exercise C

1) How can data be collected through the Observation Method?

....
....
....

2) Distinguish between the observation and the interview method of data collection.

....
....
....

3.5.3 Through Local Reporters and Correspondents

Under this method, local investigators/agents or correspondents are appointed in different parts of the area under investigation. This method is generally adopted by government departments in those cases where regular information is to be collected. This method is also useful for newspapers, magazines, radio and TV news channels. This method has been used when regular information is required

and a high degree of accuracy is not of much importance.

Merits

1) This method is cheap and economical for extensive investigations.

2) It gives results easily and promptly.

3) It can cover a wide area under investigation.

Limitations

1) The data obtained may not be reliable.

2) It gives approximate and rough results.

3) It is unsuited where a high degree of accuracy is desired.

4) As the agent/reporter or correspondent uses his own judgement, his personal bias may affect the accuracy of the information sent.

3.5.4 Questionnaire and Schedule Methods

Questionnaire and schedule methods are the popular and common methods for collecting primary data in business research. Both the methods comprise a list of questions arranged in a sequence pertaining to the investigation. Let us study these methods in detail one after another.

i) **Questionnaire Method**

Under this method, questionnaires are sent personally or by post to various informants with a request to answer the questions and return the questionnaire. If the questionnaire is posted to informants, it is called a **Mail Questionnaire**. Sometimes questionnaires may also sent through E-mail depending upon the nature of study and availability of time and resources. After receiving the questionnaires the informants read the questions and record their responses in the space meant for the purpose on the questionnaire. It is desirable to send the quetionnaire with self-addressed envelopes for quick and high rate of response.

Merits

1) You can use this method in cases where informants are spread over a vast geographical area.

2) Respondents can take their own time to answer the questions. So the researcher can obtain original data by this method.

3) This is a cheap method because its mailing cost is less than the cost of personal visits.

4) This method is free from bias of the investigator as the information is given by the respondents themselves.

5) Large samples can be covered and thus the results can be more reliable and dependable.

Limitations

1) Respondents may not return filled in questionnaires, or they can delay in replying to the questionnaires.

2) This method is useful only when the respondents are educated and co-operative.

3) Once the questionnaire has been despatched, the investigator cannot modify the questionnaire.

4) It cannot be ensured whether the respondents are truly representative.

ii) **Schedule Method**

As discussed above, a Schedule is also a list of questions, which is used to collect the data from the field. This is generally filled in by the researcher or the enumerators. If the scope of the study is wide, then the researcher appoints people who are called enumerators for the purpose of collecting the data. The enumerators go to the informants, ask them the questions from the schedule in the order they are listed and record the responses in the space meant for the answers in the schedule itself. For example, the population census all over the world is conducted through this method. The difference between questionnaire and schedule is that the former is filled in by the informants, the latter is filled in by the researcher or enumerator.

Merits

1) It is a useful method in case the informants are illiterates.

2) The researcher can overcome the problem of non-response as the enumerators go personally to obtain the information.

3) It is very useful in extensive studies and can obtain more reliable data.

Limitations

1) It is a very expensive and time-consuming method as enumerators are paid persons and also have to be trained.

2) Since the enumerator is present, the respondents may not respond to some personal questions.

3) Reliability depends upon the sincerity and commitment in data collection.

The success of data collection through the questionnaire method or schedule method depends on how the questionnaire has been designed.

Designing the Questionnaire

The success of collecting data either through the questionnaire method or through the schedule method depends largely on the proper design of the questionnaire. This is a specialised job and a requires high degree of skill, experience, thorough knowledge of the research topic, ability to frame questions and a great deal of patience. There are no hard and fast rules in designing the questionnaire. However, the following general guidelines may be helpful in this connection.

1   The number of questions should be minimised as far as possible because informants may not like to spend much time to answer a lengthy questionnaires.

1   The questions should be precise, clear and unambiguous. Lengthy questions tend to confuse the informant.

1   Choose the appropriate type of questions. Generally there are five kinds of questions used in questionnaires. They are as follows :

   i)   Simple choice questions which offer the respondents a choice between two answers, such as, 'Yes' or 'No', 'Right' or 'Wrong'. 'Do you own a computer?' Can easily be answered with 'Yes' or 'No'.

   ii) Multiple choice questions are often used as a follow-up to simple choice questions. This type of questions provide a choice between a number of factors that might influence informant preferences. For example, where do you sell your agricultural products? a) In village market, b) In a regulated market, c) To commission agent, d) Any other…

   iii) Open-ended questions allow the informants to give any related answer in their own words. For example, what should be done to enhance the practical utility of commerce programmes?

   iv) Specific questions which require specific information. For example, "From where did you take the loan for your business."

   v)   Scaled questions are used to record how strongly the opinions are expressed. For example, How do you rate the facilities provided by the market committee?

        a) Very good, b) Good, c) Normal, d) Bad, or e) Very bad.

1   The questions should be arranged in a logical sequence to avoid embarrassment. For example, asking a question how many children do you have? Then the next question : Are you married?

1   Questions which require calculations should be avoided. For example, question regarding yearly income of the respondents who are getting daily wage or piece wages, should not be asked.

Pilot testing or Pre-testing the Questionnaire

Before finalising the questionnaire, it is desirable to carry out a preliminary experiment on a sample basis. The investigator should examine each question to ensure that the question is not confusing, leading to biased responses etc. The real test of a questionnaire is how it performs under actual conditions of data collection. This test can be carried out among small groups of subjects in order to provide an estimate of the time needed for responding to the survey. The questionnaire pre-test serves the same role in questionnaire design as testing a new product in the market. As test marketing provides the real test of customer reactions to the product and the accompanying marketing programmes, in the same way, the pre-test provides the real test of the questionnaire. Therefore this work must be done with utmost care and caution to yield good results.

Specimen Questionnaire

The following specimen questionnaire incorporates most of the qualities which we have discussed above. It relates to 'Computer User Survey'.

Computer User Survey

1. What brand of Computer do you primarily use?

   (i)   IBM   (ii)   Compaq
   (iii) HCL   (iv) Dell
   (v) Siemens   (vi) Any other
   _____
   (please specify)

2. Where was the computer purchased?
   (i)   Computer store   (ii)   Mail order
   (iii) Manufacturer   (iv) Company Dealer
   (v) Any other _____

3. How long have you been using computers? _____years
   _____months.

4. In a week about how many hours do you spend on the computer _____
   hours?

5. Which database management package do you use most often?
   (i)   Dbase-II   (ii)   Dbase-III
   (iii) Lotus 1,2,3   (iv) MS-Excel
   (v) Oracle   (vi) Any other
   _____
   (please specify)

6. Does the computer, that you primarily use, have a hard disk
   Yes ☐   No ☐

7. Where did you obtain the software that you use?
   (i)   Computer user group   (ii)   Regular dealer
   (iii) Mail order   (iv) Directly from Software
   (v) Any other _____   dealer

8. On the following 9-point scale, rate the degree of difficulty that you have
   encountered in using the computer.

   Extremely difficult ⟶ 1 2 3 4 5 6 7 8 9 ⟵ Not difficult

9. If you have to purchase a personal computer today, which one would
   you be most likely to purchase?

   (i)   IBM   (ii)   Compaq
   (iii) HCL   (iv) Dell
   (v) Siemens   (vi) Any other
   _____
   (please specify)

10. What is your sex   ☐ Male   ☐ Female

11. Please state your date of birh ....
                          Month   Day   Year

12. Your Qualifications
    (i)   Secondary   (ii)   Sr. Secondary
    (iii) Graduate   (iv) Post-graduate
    (v) Doctorate   (vi) Any other
    _____
    (please specify)

13. Which of the following best describe you primary field of
employment.
    (i)   Medical   (ii)   Education
    (iii) Business   (iv) Government
    (v) Technical   (vi) Any other _____
                                     (please specify)

14. What is your current Salary?

3.6 CHOICE OF SUITABLE METHOD

You have noticed that there are various methods and techniques for the collection of primary data. You should be careful while selecting the method which should be appropriate and effective. The selection of the methods depends upon various factors like scope and objectives of the inquiry, time, availability of funds, subject matter of the research, the kind of information required, degree of accuracy etc. As apprised, every method has its own merits and demerits. For example, the observation method is suitable for field surveys when the incident is really happening, the interview method is suitable where direct observation is not possible. Local reporter/correspondent method is suitable when information is required at regular intervals. The questionnaire method is appropriate in extensive enquiries where sample is large and scattered over large geographical areas and the respondents are able to express their responses in writing. The Schedule method is suitable in case respondents are illiterate.

Self Assessment Exercise D

1) List out the methods of collecting primary data.

    ....
    ....
    ....

2) Point out the major problems in constructing questionnaires.

    ....
    ....
    ....

3) Distinguish between direct personal interview and indirect interview. Give suitable examples.

    ....
    ....
    ....

4) Distinguish between Schedule and Questionnaire?

    ....
    ....
    ....

5) Are the following statement true or false?

    a) Interview method introduces more bias than the use of questionnaire.

    b) 'Yes' or 'No' type questions should not be used in questionnaires unless only one of the two answer is possible.

    c) Open questions are more difficult than most other types to tabulate.

3.7 LET US SUM UP

In this unit we elaborated on the meaning of data, methods of data collection, merits and limitations of data collection, precautions which are needed for the collection of data.

The information collected from various processes for a specific purpose is called data. Statistical data may be either primary data or secondary data. Data which is collected originally for a specific purpose is called primary data. The data which is already collected and processed by some one else and is being used now in the present study, is called secondary data. Secondary data can be obtained either from published sources or unpublished sources. It should be used if it is reliable, suitable and adequate, otherwise it may result in misleading conclusions. It has its own merits and demerits. There are several problems in the collection of primary data. These are: tools and techniques of data collection, degree of accuracy, designing the questionnaire, selection and training of enumerators, problem of tackling non-responses and other administrative aspects.

Several methods are used for collection of primary data. These are: observation, interview, questionnaire and schedule methods. Every method has its own merits and demerits. Hence, no method is suitable in all situations. The suitable method can be selected as per the needs of the investigator which depends on objective nature and scope of the enquiry, availability of funds and time.

3.8   KEY WORDS

Data: Quantitative or/ and qualitative information, collected for study and analysis.

Interview: A method of collecting primary data by meeting the informants and asking the questions.

Observation: The process of observing individuals in controlled situations.

Questionnaire: is a device for collection of primary data containing a list of questions pertaining to enquiry, sent to the informants, and the informant himself writes the answers.

Primary Data: Data that is collected originally for the first time.

Secondary Data: Data which were collected and processed by someone else but are being used in the present study.

Published Sources: Sources which consist of published statistical information.

Schedule: is a device for collection of primary data containing a list of questions to be filled in by the enumerators who are specially appointed for that purpose.

3.9   ANSWERS TO SELF ASSESSMENT EXERCISES

A. 1)   In many business companies, some of the data required for statistical analysis are obtained from internal sources like computer files of accounting data. Together with internal data, business often uses data from external sources. For example, aggregate data on national economic activity are readily available from CSO, annual report of Ministry of Labour, Government of India.

   2)   Data which is collected originally is called primary data and the same collected by others are called secondary data. For example, a researcher interested in knowing what consumer's choice about the brand of

toothpaste, he or she must make a survey and collect data on the opinions of the consumer. This is called primary data. The data obtained from published and unpublished sources is called secondary data.

B. 1)   ι   http://www.bis.org.in

       ι   http://www.business-today.com

       ι   http://www.businessonlineindia.com

       ι   http://www.indiacofee.org

       ι   http://www.dgft.nic.in

       IV.   i) No,   ii) No

D. 1)   Observation Method, interview method, data collection through reporters/correspondents, questionnaries and schedule methods.

   2)   a)   What information will be sought?

         b)   What type of questionnaire will be required?

         c)   How many questions will be used?

         d)   What will the content of the individual questions be?

         e)   How will those questionnaires be administered?

   3)   i)   Personal Interview: Under this method, the investigator collects information personally from the concerned sources.

         ii)   Indirect Interview : Under this method, the investigator contacts third parties or witnesses capable of supplying the necessary information.

   5)   i)   True   ii) True   iii) True.

3.10   TERMINAL QUESTIONS

1)   What precautions would you take while using the data from secondary sources.

2)   Explain what precautions must be taken while designing a questionnaire in order that it may be really useful. Illustrate your answer giving suitable examples.

3)   Construct a suitable questionnaire containing not more than twenty five questions pertaining to the sales promotion of your company's product.

4)   Distinguish between the following :
     a) Primary and Secondary Data
     b) Internal and External Data
     c) A Schedule and Questionnaire

5)   Explain the various methods of collecting primary data pointing out their merits and demerits?

6)   What is the need for pre-testing the drafted questionnaire.

> **Note:** These questions/exercises will help you to understand the unit better. Try to write answers for them. But do not submit your answers to the university for assessment. These are for your practice only.

3.11 FURTHER READING

The following text books may be used for more indepth study on the topics dealt with in this unit.

Kothari, C.R. 2004. *Research Methodology Methods and Techniques,* New Age International (P) Limited : New Delhi.

Rao K.V. 1993. *Research Methodology in Commerce and Management,* Sterling Publishers Private Limited : New Delhi.

Sadhu, A.N. and A. Singh, 1980. *Research Methodology in Social Sciences,* Sterling Publishers Private Limited : New Delhi.

UNIT 4   SAMPLING

STRUCTURE

4.0   OBJECTIVES

After studying this Unit, you should be able to:

l   distinguish between census and sampling study,

l   explain various reasons for opting for the sample method,

l   explain the different methods of sampling and their advantages and disadvantages,

l   describe the sampling and non-sampling errors and minimize them, and

l   design a representative sample from a population keeping both cost and precision in mind.

4.1   INTRODUCTION

In the previous Unit 3, we have studied the types of data (primary and secondary data) and various methods and techniques of collecting the primary data. The desired data may be collected by selecting either census method or sampling method.

Researchers usually cannot make direct observations of every unit of the population they are studying for a variety of reasons. Instead, they collect data from a subset of population – a sample – and use these observations drawn to make inferences about the entire population. Ideally, the characteristics of a sample should correspond to the characteristics of a population from which the sample was drawn. In that case, the conclusions drawn from a sample are probably applicable to the entire population.

In this Unit, we shall discuss the basics of sampling, particularly how to get a sample that is representative of a population. It covers different methods of drawing samples which can save a lot of time, money and manpower in a

variety of situations. These include random sampling methods, such as, simple random sampling, stratified sampling, systematic sampling, multistage sampling, cluster sampling methods (and non-random sampling methods viz., convenience sampling, judgement sampling and quota sampling. The advantages and disadvantages of sampling and census are covered. How to determine the sample size of a given population is also discussed.

4.2   CENSUS AND SAMPLE

Let us try to understand the terms 'census' and 'sample' with the help of an illustration. Suppose you wish to study the 'impact of T.V. advertisements on children in Delhi, then you have to collect relevant information from the children residing in Delhi who view T.V. Alternatively, we can say this is the population (statistical terminology) for your study. If you collect the data from all of them not leaving a single child, it known as Census method of data collection. This means studying the whole population. Otherwise, if you select only some children from among them for gathering the desired information for the study, because it is not feasible to gather the information from all the children, then it is known as **Sample** for data collection. Therefore, a sample is a subset of a statistical population whose characteristics are studied to know the information about the whole population. When dealing with people, it can be defined as a set of respondents (people) selected from a population for the purpose of a survey. A population is a group of individual persons, objects, items or any other units from which samples are taken for measurement.

The numerical characteristics of a population are called **parameters**. They are fixed and usually of unknown quantity. For example, the average (μ) height of all Indian male adults is a population parameter. The numerical characteristics of the sample data such as the mean, variance or proportion are called **sample statistics**. It can be used to provide estimates of the corresponding population parameters. For example, the average (x) height of a sample of 1000 Indian male adults residing in Delhi is a sample statistic. The process of selecting a representative sample for the purpose of inferring the characteristics of population is called sampling.

Webster defines a survey as 'the action of ascertaining facts regarding conditions or the condition of something to provide *exact information* especially to persons responsible or interested' and as 'a systematic collection and analysis of data on some aspect of an area or group.' Unless the researcher makes a systematic collection of data followed by careful analysis and interpretation of data, the data cannot become exact information. Surveys can be divided into two categories on the basis of their extensiveness, namely, census and sample survey. A complete survey of population is called a census. It involves covering all respondents, items, or units of the population. For example, if we want to know the wage structure of the textile industry in the country, then one approach is to collect the data on the wages of each and every worker in the textile industry. On the other hand, a sample is a representative subset of population. Thus in a sample survey we cover only a sample of respondents, items or units of population we are interested in and then draw inferences about the whole population.

The following are the advantages of census:

1) In a census each and every respondent of the population is considered and various population parameters are compiled for information.

2) The information obtained on the basis of census data is more reliable and accurate. It is an adopted method of collecting data on exceptional matters like child labour, distribution by sex, educational level of the people etc.

3) If we are conducting a survey for the first time we can have a census instead of sample survey. The information based on this census method becomes a base for future studies. Similarly, some of the studies of special importance like population data are obtained only through census.

4.3 WHY SAMPLING?

One of the decisions to be made by a researcher in conducting a survey is whether to go for a census or a sample survey. We obtain a sample rather than a complete enumeration (a census) of the population for many reasons. The most important considerations for this are: cost, size of the population, accuracy of data, accessibility of population, timeliness, and destructive observations.

1) **Cost:** The cost of conducting surveys through census method would be prohibitive and sampling helps in substantial cost reduction of surveys. Since most often the financial resources available to conduct a survey are scarce, it is imperative to go for a sample survey than census.

2) **Size of the Population:** If the size of the population is very large it is difficult to conduct a census if not impossible. In such situations sample survey is the only way to analyse the characteristics of a population.

3) **Accuracy of Data:** Although reliable information can be obtained through census, sometime the accuracy of information may be lost because of a large population. Sampling involves a small part of the population and a few trained people can be involved to collect accurate data. On the other hand, a lot of people are required to enumerate all the observations. Often it becomes difficult to involve trained manpower in large numbers to collect the data thereby compromising accuracy of data collected. In such a situation a sample may be more accurate than a census. A sloppily conducted census can provide less reliable information than a carefully obtained sample.

4) **Accessibility of Population:** There are some populations that are so difficult to get access to that only a sample can be used, e.g., people in prison, birds migrating from one place to another place etc. The inaccessibility may be economic or time related. In a particular study, population may be so costly to reach, like the population of planets, that only a sample can be used.

5) **Timeliness:** Since we are covering a small portion of a large population through sampling, it is possible to collect the data in far less time than covering the entire population. Not only does it take less time to collect the data through sampling but the data processing and analysis also takes less time because fewer observations need to be covered. Suppose a company wants to get a quick feedback from its consumers on assessing their perceptions about a new improved detergent in comparison to an existing version of the detergent. Here the time factor is very significant. In such situations it is better to go for a sample survey rather than census because it reduces a lot of time and product launch decision can be taken quickly.

6) **Destructive Observations:** Sometimes the very act of observing the desired characteristics of a unit of the population destroys it for the intended

use. Good examples of this occur in quality control. For example, to test the quality of a bulb, to determine whether it is defective, it must be destroyed. To obtain a census of the quality of a lorry load of bulbs, you have to destroy all of them. This is contrary to the purpose served by quality-control testing. In this case, only a sample should be used to assess the quality of the bulbs. Another example is blood test of a patient.

The disadvantages of sampling are few but the researcher must be cautious. These are risk, lack of representativeness and insufficient sample size each of which can cause errors. If researcher don't pay attention to these flaws it may invalidate the results.

1) **Risk:** Using a sample from a population and drawing inferences about the entire population involves risk. In other words the risk results from dealing with a part of a population. If the risk is not acceptable in seeking a solution to a problem then a census must be conducted.

2) **Lack of representativeness:** Determining the representativeness of the sample is the researcher's greatest problem. By definition, 'sample' means a representative part of an entire population. It is necessary to obtain a sample that meets the requirement of representativeness otherwise the sample will be biased. The inferences drawn from nonreprentative samples will be misleading and potentially dangerous.

3) **Insufficient sample size:** The other significant problem in sampling is to determine the size of the sample. The size of the sample for a valid sample depends on several factors such as extent of risk that the researcher is willing to accept and the characteristics of the population itself.

4.4   ESSENTIALS OF A GOOD SAMPLE

It is important that the sampling results must reflect the characteristics of the population. Therefore, while selecting the sample from the population under investigation it should be ensured that the sample has the following characteristics:

1) A sample must represent a true picture of the population from which it is drawn.

2) A sample must be unbiased by the sampling procedure.

3) A sample must be taken at random so that every member of the population of data has an equal chance of selection.

4) A sample must be sufficiently large but as economical as possible.

5) A sample must be accurate and complete. It should not leave any information incomplete and should include all the respondents, units or items included in the sample.

6) Adequate sample size must be taken considering the degree of precision required in the results of inquiry.

Self Assessment Exercise A

1) What do you mean by census and sample methods for data collection?

....

....

....

....

2) Explain whether census or sample is more appropriate in the following situations?

    a) To test the quality of a soft drink.

    ....

    ....

    b) To enumerate eligible voters of an assembly constituency.

    ....

    ....

    c) To know the opinion of consumers on launching a new product.

    ....

    ....

3) Fill in the blanks

    a) If the sample does not represent the population characteristics, we call it a ——————— sample.

    b) One of the major advantages of sampling is that it helps in ——— reduction.

    c) A sample must be ——— large but as ——— as possible.

4.5   METHODS OF SAMPLING

If money, time, trained manpower and other resources were not a concern, the researcher could get most accurate data from surveying the entire population of interest. Since most often the resources are scarce, the researcher is forced to go for sampling. But the real purpose of the survey is to know the characteristics of the population. Then the question is with what level of confidence will the researcher be able to say that the characteristics of a sample represent the entire population. Using a combination of tasks of hypotheses and unbiased sampling methods, the researcher can collect data that actually represents the characteristics of the entire population from which the sample was taken. To ensure a high level of confidence that the sample represents the population it is necessary that the sample is unbiased and sufficiently large.

It was scientifically proved that if we increase the sample size we shall be that much closer to the characteristics of the population. Ultimately, if we cover each and every unit of the population, the characteristics of the sample will be equal to the characteristics of the population. That is why in a census there is no sampling error. Thus, "generally speaking, the larger the sample size, the less sampling error we have."

The statistical meaning of bias is error. The sample must be error free to make it an unbiased sample. In practice, it is impossible to achieve an error free sample even using unbiased sampling methods. However, we can minimize the error by employing appropriate sampling methods.

The various sampling methods can be classified into two categories. These are random sampling methods and non-random sampling methods. Let us discuss them in detail.

4.5.1 Random Sampling Methods

The random sampling method is also often called probability sampling. In random sampling all units or items in the population have a chance of being chosen in the sample. In other words a random sample is a sample in which each element of the population has a known and non-zero chance of being selected. Random sampling always produces the smallest possible sampling error. In the real sense, the size of the sampling error in a random sample is affected only by a random chance. Because a random sample contains the least amount of sampling error, we may say that it is an unbiased sample. Remember that we are not saying that a random sample contains no error, but rather the minimum possible amount of error. The major advantage of random sampling is that it is possible to quantify the magnitude of the likely error in the inference made and this will help in building confidence in drawing inferences.

The following are the important methods of random sampling:

1) Simple Random Sampling

2) Systematic Sampling

3) Stratified Random Sampling

4) Cluster Sampling

5) Multistage Sampling

1. Simple Random Sampling: The most commonly used random sampling method is simple random sampling method. A simple random sample is one in which each item in the total population has an equal chance of being included in the sample. In addition, the selection of one item for inclusion in the sample should in no way influence the selection of another item. Simple random sampling should be used with a homogeneous population, that is, a population consisting of items that possess the same attributes that the researcher is interested in. The characteristics of homogeneity may include such as age, sex, income, social/religious/political affiliation, geographical region etc.

The best way to choose a simple random sample is to use random number table. A random sampling method should meet the following criteria.

a) Every member of the population must have an equal chance of inclusion in the sample.

b) The selection of one member is not affected by the selection of previous members.

The random numbers are a collection of digits generated through a probabilistic mechanism. The random numbers have the following properties:

i) The probability that each digit (0,1,2,3,4,5,6,7,8,or 9) will appear at any place is the same. That is 1/10.

ii) The occurrence of any two digits in any two places is independent of each other.

Each member of a population is assigned a unique number. The members of the population chosen for the sample will be those whose numbers are identical to the ones extracted from the random number table in succession until the desired sample size is reached. An example of a random number table is given below.

Table 1: Table of Random Numbers

	1	2	3	4	5	6	7	8	9	10
1	96268	11860	83699	38631	90045	69696	48572	05917	51905	10052
2	03550	59144	59468	37984	77892	89766	86489	46619	50236	91136
3	22188	81205	99699	84260	19693	36701	43233	62719	53117	71153
4	63759	61429	14043	44095	84746	22018	19014	76781	61086	90216
5	55006	17765	15013	77707	54317	48862	53823	52905	70754	68212
6	81972	45644	12600	01951	72166	52682	37598	11955	73018	23528
7	06344	50136	33122	31794	86723	58037	36065	32190	31367	96007
8	92363	99784	94169	03652	80824	33407	40837	97749	18361	72666
9	96083	16943	89916	55159	62184	86206	09764	20244	88388	98675
10	92993	10747	08985	44999	35785	65036	05933	77378	92339	96151
11	95083	70292	50394	61947	65591	09774	16216	63561	59751	78771
12	77308	60721	96057	86031	83148	34970	30892	53489	44999	18021
13	11913	49624	28519	27311	61586	28576	43092	69971	44220	80410
14	70648	47484	05095	92335	55299	27161	64486	71307	85883	69610
15	92771	99203	37786	81142	44271	36433	31726	74879	89384	76886
16	78816	20975	13043	55921	82774	62745	48338	88348	61211	88074
17	79934	35392	56097	87613	94627	63622	08110	16611	88599	02890
18	64698	83376	87527	36897	17215	74339	69856	43622	22567	11518
19	44212	12995	03581	37618	94851	63020	65348	55857	91742	79508
20	89292	00204	00579	70630	37136	50922	83387	15014	51838	81760
21	08692	87237	87879	01629	72184	33853	95144	67943	19345	03469
22	67927	76855	50702	78555	97442	78809	40575	79714	06201	34576
23	62167	94213	52971	85794	68067	78814	40103	70759	92129	46716
24	45828	45441	74220	84157	23241	49332	23646	09390	13031	51569
25	01164	35307	26526	80335	58090	85871	07205	31749	40571	51755
26	29283	31581	04359	45538	41435	61103	32428	94042	39971	63678
27	19868	49978	81699	84904	50163	22652	07845	71308	00859	87984
28	14292	93587	55960	23159	07370	65065	06580	46285	07884	83928
29	77410	52135	29495	23032	83242	89938	40516	27252	55565	64714
30	36580	06921	35675	81645	60479	71035	99380	59759	42161	93440
31	07780	18093	31258	78156	07871	20369	53977	08534	39433	57216
32	07548	08454	36674	46255	80541	42903	37366	21164	97516	66181
33	22023	60448	69344	44260	90570	01632	21002	24413	04671	05665
34	20827	37210	57797	34660	32510	71558	78228	42304	77197	79168
35	47802	79270	48805	59480	88092	11441	96016	76091	51823	94442
36	76730	86591	18978	25479	77684	88439	34112	26052	57112	91653
37	26439	02903	20935	76297	15290	84688	74002	09467	41111	19194
38	32927	83426	07848	59372	44422	53372	27823	25417	27150	21750
39	51484	05286	77103	47284	00578	88774	15293	50740	07932	87633
40	45142	96804	92834	26886	70002	96643	36008	02239	93563	66429

To select a random sample using simple random sampling method we should follow the steps given below:

i)   Determine the population size (N).

ii) Determine the sample size (n).

iii) Number each member of the population under investigation in serial order. Suppose there are 100 members number them from 00 to 99.

iv) Determine the starting point of selecting sample by randomly picking up a page from random number tables and dropping your finger on the page blindly.

v)   Choose the direction in which you want to read the numbers (from left to right, or right to left, or down or up).

vi) Select the first 'n' numbers whose X digits are between 0 and N. If N = 100 then X would be 2, if N is a four digit number then X would be 3 and so on.

vii) Once a number is chosen, do not use it again.

viii) If you reach the end point of the table before obtaining 'n' numbers, pick another starting point and read in a different direction and then use the first X digit instead of the last X digits and continue until the desired sample is selected.

Example: Suppose you have a list of 80 students and want to select a sample of 20 students using simple random sampling method. First assign each student a number from 00 to 79. To draw a sample of 20 students using random number table, you need to find 20 two-digit numbers in the range 00 to 79. You can begin any where and go in any direction. For example, start from the 6th row and 1st column of the random number table given in this Unit. Read the last two digits of the numbers. If the number is within the range (00 to 79) include the number in the sample. Otherwise skip the number and read the next number in some identified direction. If a number is already selected omit it. In the example starting from 6th row and 1st column and moving from left to right direction the following numbers are considered to selected 20 numbers for sample.

81**72** 45**44** 126**00** 019**51** 721**66** 52682 37598 11**55** 730**18** 235**28**

06344 501**36** 33122 31794 867**23** 580**37** 360**65** 32190 313**67** 960**07**

923**63** 99784 941**69** 036**52** 808**24** 33407 40837 977**49** 18361 72666

The bold faced digits in the one's and ten's place value indicate the selected numbers for the sample. Therefore, the following are the 20 numbers chosen as sample.

72	44	00	51	66	55	18	28
36	22	23	37	65	67	07	63
69	52	24	49				

Advantages

i)   The simple random sample requires less knowledge about the characteristics of the population.

ii) Since sample is selected at random giving each member of the population equal chance of being selected the sample can be called as unbiased sample. Bias due to human preferences and influences is eliminated.

iii) Assessment of the accuracy of the results is possible by sample error estimation.

iv) It is a simple and practical sampling method provided population size is not large.

Limitations

i)   If the population size is large, a great deal of time must be spent listing and numbering the members of the population.

ii) A simple random sample will not adequately represent many population characteristics unless the sample is very large. That is, if the researcher is interested in choosing a sample on the basis of the distribution in the population of gender, age, social status, a simple random sample needs to be very large to ensure all these distributions are representative of the population. To obtain a representative sample across multiple population attributes we should use stratified random sampling.

2. **Systematic Sampling:** In systematic sampling the sample units are selected from the population at equal intervals in terms of time, space or order. The selection of a sample using systematic sampling method is very simple. From a population of 'N' units, a sample of 'n' units may be selected by following the steps given below:

   i)   Arrange all the units in the population in an order by giving serial numbers from 1 to N.

   ii) Determine the sampling interval by dividing the population by the sample size. That is, K=N/n.

   iii) Select the first sample unit at random from the first sampling interval (1 to K).

   iv) Select the subsequent sample units at equal regular intervals.

For example, we want to have a sample of 100 units from a population of 1000 units. First arrange the population units in some serial order by giving numbers from 1 to 1000. The sample interval size is K=1000/100=10. Select the first sample unit at random from the first 10 units (i.e. from 1 to 10). Suppose the first sample unit selected is 5, then the subsequent sample units are 15, 25, 35,.........995. Thus, in the systematic sampling the first sample unit is selected at random and this sample unit in turn determines the subsequent sample units that are to be selected.

Advantages

i)   The main advantage of using systematic sample is that it is more expeditious to collect a sample systematically since the time taken and work involved is less than in simple random sampling. For example, it is frequently used in exit polls and store consumers.

ii) This method can be used even when no formal list of the population units is available. For example, suppose if we are interested in knowing the opinion of consumers on improving the services offered by a store we may simply choose

every kth (say 6th) consumer visiting a store provided that we know how many consumers are visiting the store daily (say 1000 consumers visit and we want to have 100 consumers as sample size).

Limitations

i) If there is periodicity in the occurrence of elements of a population, the selection of sample using systematic sample could give a highly un-representative sample. For example, suppose the sales of a consumer store are arranged chronologically and using systematic sampling we select sample for 1st of every month. The 1st day of a month can not be a representative sample for the whole month. Thus in systematic sampling there is a danger of order bias.

ii) Every unit of the population does not have an equal chance of being selected and the selection of units for the sample depends on the initial unit selection. Regardless how we select the first unit of sample, subsequent units are automatically determined lacking complete randomness.

3. **Stratified Random Sampling:** The stratified sampling method is used when the population is heterogeneous rather than homogeneous. A heterogeneous population is composed of unlike elements such as male/female, rural/urban, literate/illiterate, high income/low income groups, etc. In such cases, use of simple random sampling may not always provide a representative sample of the population. In stratified sampling, we divide the population into relatively homogenous groups called strata. Then we select a sample using simple random sampling from each stratum. There are two approaches to decide the sample size from each stratum, namely, proportional stratified sample and disproportional stratified sample. With either approach, the stratified sampling guarantees that every unit in the population has a chance of being selected. We will now discuss these two approaches of selecting samples.

i) *Proportional Stratified Sample:* If the number of sampling units drawn from each stratum is in proportion to the corresponding stratum population size, we say the sample is proportional stratified sample. For example, let us say we want to draw a stratified random sample from a heterogeneous population (on some characteristics) consisting of rural/urban and male/female respondents. So we have to create 4 homogeneous sub groups called stratums as follows:

Urban		Rural	
Male	Female	Male	Female

To ensure each stratum in the sample will represent the corresponding stratum in the population we must ensure each stratum in the sample is represented in the same proportion to the stratums as they are in the population. Let us assume that we know (or can estimate) the population distribution as follows: 65% male, 35% female and 30% urban and 70% rural. Now we can determine the approximate proportions of our 4 stratums in the population as shown below.

Urban		Rural	
Male	Female	Male	Female
$0.30 \times 0.65 = 0.195$	$0.30 \times 0.35 = 0.105$	$0.70 \times 0.65 = 0.455$	$0.70 \times 0.35 = 0.245$

Thus a representative sample would be composed of 19.5% urban-males, 10.5% urban-females, 45.5% rural-males and 24.5% rural females. Each percentage should be multiplied by the total sample size needed to arrive at the actual

sample size required from each stratum. Suppose we require 1000 samples then the required sample in each stratum is as follows:

Urban-male	0.195×1000	$= 195$
Urban-female	0.105×1000	$= 105$
Rural-male	0.455×1000	$= 455$
Rural-female	0.245×1000	$= 245$
Total:		1,000

ii) *Disproportional Stratified Sample:* In a disproportional stratified sample, sample size for each stratum is not allocated on a proportional basis with the population size, but by analytical considerations of the researcher such as stratum variance, stratum population, time and financial constraints etc. For example, if the researcher is interested in finding differences among different stratums, disproportional sampling should be used. Consider the example of income distribution of households. There is a small percentage of households within the high income brackets and a large percentage of households within the low income brackets. The income among higher income group households has higher variance than the variance among the lower income group households. To avoid under-representation of higher income groups in the sample, a disproportional sample is taken. This indicates that as the variability within the stratum increases sample size must increase to provide accurate estimates and vice-versa.

Suppose in our example of urban/rural and male/female stratum populations, the stratum estimated variances (s^2) are as follows. However, the variance is discussed in Unit 9 of this course.

Urban-male 3.0; Urban-female 5.5; Rural-males 2.5; Rural-females 1.75.

The above figures are, normally, estimated on the basis of previous knowledge of a researcher.

Then the allocation of sample size of 1000 for each strata using disproportional stratified sampling method will be as shown in the following table:

Stratum	Stratum population proportion (P_i)	Stratum variance (σ_i^2)	Stratum standard deviation (σ_i)	$P_i \times \sigma_i$	Sample size $(P_i \times \sigma_i \times 1000)/ \Sigma \ P_i \sigma_i)$
Urban-male	0.195	3.0	1.73	0.338	207
Urban-female	0.105	5.5	2.35	0.246	151
Rural-male	0.455	2.5	1.58	0.719	442
Rural-female	0.245	1.75	1.32	0.324	199
			Total	1.628	1000

Advantages

a) Since the sample are drawn from each of the stratums of the population, stratified sampling is more representative and thus more accurately reflects characteristics of the population from which they are chosen.

b) It is more precise and to a great extent avoids bias.

c) Since sample size can be less in this method, it saves a lot of time, money and other resources for data collection.

Limitations

a) Stratified sampling requires a detailed knowledge of the distribution of attributes or characteristics of interest in the population to determine the homogeneous groups that lie within it. If we cannot accurately identify the homogeneous groups, it is better to use simple random sample since improper stratification can lead to serious errors.

b) Preparing a stratified list is a difficult task as the lists may not be readily available.

4. Cluster Sampling: In cluster sampling we divide the population into groups having heterogenous characteristics called clusters and then select a sample of clusters using simple random sampling. We assume that each of the clusters is representative of the population as a whole. This sampling is widely used for geographical studies of many issues. For example if we are interested in finding the consumers' (residing in Delhi) attitudes towards a new product of a company, the whole city of Delhi can be divided into 20 blocks. We assume that each of these blocks will represent the attitudes of consumers of Delhi as a whole, we might use cluster sampling treating each block as a cluster. We will then select a sample of 2 or 3 clusters and obtain the information from consumers covering all of them. The principles that are basic to the cluster sampling are as follows:

i) The differences or variability within a cluster should be as large as possible. As far as possible the variability within each cluster should be the same as that of the population.

ii) The variability between clusters should be as small as possible. Once the clusters are selected, all the units in the selected clusters are covered for obtaining data.

Advantages

a) The cluster sampling provides significant gains in data collection costs, since traveling costs are smaller.

b) Since the researcher need not cover all the clusters and only a sample of clusters are covered, it becomes a more practical method which facilitates fieldwork.

Limitations

a) The cluster sampling method is less precise than sampling of units from the whole population since the latter is expected to provide a better cross-section of the population than the former, due to the usual tendency of units in a cluster to be homogeneous.

b) The sampling efficiency of cluster sampling is likely to decrease with the decrease in cluster size or increase in number of clusters.

The above advantages or limitations of cluster sampling suggest that, in practical situations where sampling efficiency is less important but the cost is of greater significance, the cluster sampling method is extensively used. If the division of clusters is based on the geographic sub-divisions, it is known as area sampling. In cluster sampling instead of covering all the units in each cluster we can resort to sub-sampling as two-stage sampling. Here, the clusters are termed as primary units and the units within the selected clusters are taken as secondary units.

5. **Multistage Sampling:** We have already covered two stage sampling. Multi stage sampling is a generalisation of two stage sampling. As the name suggests, multi stage sampling is carried out in different stages. In each stage progressively smaller (population) geographic areas will be randomly selected.

A political pollster interested in assembly elections in Uttar Pradesh may first divide the state into different assembly units and a sample of assembly constituencies may be selected in the first stage. In the second stage, each of the sampled assembly constituents are divided into a number of segments and a second stage sampled assembly segments may be selected. In the third stage within each sampled assembly segment either all the house-holds or a sample random of households would be interviewed. In this sampling method, it is possible to take as many stages as are necessary to achieve a representative sample. Each stage results in a reduction of sample size.

In a multi stage sampling at each stage of sampling a suitable method of sampling is used. More number of stages are used to arrive at a sample of desired sampling units.

Advantages

a) Multistage sampling provides cost gains by reducing the data collection on costs.

b) Multistage sampling is more flexible and allows us to use different sampling procedures in different stages of sampling.

c) If the population is spread over a very wide geographical area, multistage sampling is the only sampling method available in a number of practical situations.

Limitations

a) If the sampling units selected at different stages are not representative multistage sampling becomes less precise and efficient.

4.5.2 Non-Random Sampling Methods

The non-random sampling methods are also often called non-probability sampling methods. In a non-random sampling method the probability of any particular unit of the population being chosen is unknown. Here the method of selection of sampling units is quite arbitrary as the researchers rely heavily on personal judgment. Non-random sampling methods usually do not produce samples that are representative of the general population from which they are drawn. The greatest error occurs when the researcher attempts to generalise the results on the basis of a sample to the entire population. Such an error is insidious because it is not at all obvious from merely looking at the data, or even from looking at the sample. The easiest way to recognise whether a sample is representative or not is to determine whether the sample is selected randomly or not. Nevertheless, there are occasions where non-random samples are best suited for the researcher's purpose. The various non-random sampling methods commonly used are:

1) Convenience Sampling;

2) Judgement Sampling; and

3) Quota Sampling.

Let us discuss these methods in detail.

1) **Convenience Sampling:** Convenience sampling refers to the method of obtaining a sample that is most conveniently available to the researcher. For example, if we are interested in finding the overtime wage paid to employees working in call centres, it may be convenient and economical to sample

employees of call centres in a nearby area. Also, on various issues of public interest like budget, election, price rise etc., the television channels often present on-the-street interviews with people to reflect public opinion. It may be cautioned that the generalisation of results based on convenience sampling beyond that particular sample may not be appropriate. Convenience samples are best used for exploratory research when additional research will be subsequently conducted with a random sample. Convenience sampling is also useful in testing the questionnaires designed on a pilot basis. Convenience sampling is extensively used in marketing studies.

2) **Judgement Sampling:** Judgement sampling method is also known as purposive sampling. In this method of sampling the selection of sample is based on the researcher's judgment about some appropriate characteristic required of the sample units. For example, the calculation of consumer price index is based on a judgment sample of a basket of consumer items, and other related commodities and services which are expected to reflect a representative sample of items consumed by the people. The prices of these items are collected from selected cities which are viewed as typical cities with demographic profiles matching the national profile. In business judgment sampling is often used to measure the performance of salesmen/saleswomen. The salesmen/saleswomen are grouped into high, medium or low performers based on certain specified qualities. Then the sales manager may actually classify the salesmen/saleswomen working under him/her who in his/her opinion will fall in which group. Judgment sampling is also often used in forecasting election results. We may often wonder how a pollster can predict an election based on only 2% to 3% of votes covered. It is needless to say the method is biased and does not have any scientific basis. However, in the absence of any representative data, one may resort to this kind of non-random sampling.

3) **Quota Sampling:** The quota sampling method is commonly used in marketing research studies. The samples are selected on the basis of some parameters such as age, sex, geographical region, education, income, occupation etc, in order to make them as representative samples. The investigators, then, are assigned fixed quotas of the sample meeting these population characteristics. The purpose of quota sampling is to ensure that various sub-groups of the population are represented on pertinent sample characteristics to the extent that the investigator desires. The stratified random sampling also has this objective but should not be confused with quota sampling. In the stratified sampling method the researcher selects a random sample from each group of the population, where as, in quota sampling, the interviewer has a quota fixed for him/her to achieve. For example, if a city has 10 market centres, a soft drink company may decide to interview 50 consumers from each of these 10 market centres to elicit information on their products. It is entirely left to the investigator whom he/she will interview at each of the market centres and the time of interview. The interview may take place in the morning, mid day, or evening or it may be in the winter or summer.

Quota sampling has the advantage that the sample confirms the selected characteristics of the population that the researcher desires. Also, the cost and time involved in collecting the data are also greatly reduced. However, quota sampling has many limitations, as given below:

a) In quota sampling the respondents are selected according to the convenience of the field investigator rather than on a random basis. This kind of selection of sample may be biased. Suppose in our example of soft drinks, after the sample is taken it was found that most of the respondents belong to the lower income group then the purpose of conducting the survey becomes useless and the results may not reflect the actual situation.

b) If the number of parameters, on which basis the quotas are fixed, are larger then it becomes difficult for the researcher to fix the quota for each sub-group.

c) The field workers have the tendency to cover the quota by going to those places where the respondents may be willing to provide information and avoid those with unwilling respondents. For example, the investigators may avoid places where high income group respondents stay and cover only low income group areas.

Self Assessment Exercise B

1) Suppose there are 900 families residing in a colony. You are asked to select a sample of families using simple random sampling for knowing the average income. The families are identified with serial numbers 001 to 900.

   i) Select a random sample using the following random table.

29283	31581	04359	45538	41435	61103	32428	94042	39971	63678
19868	49978	81699	84904	50163	22652	07845	71308	00859	87984
14292	93587	55960	23159	07370	65065	06580	46285	07884	83928
77410	52135	29495	23032	83242	89938	40516	27252	55565	64714
36580	06921	35675	81645	60479	71035	99380	59759	42161	93440

   ii) While selecting the random sample in the above example, what are the random numbers you have rejected and why?

   ....

   ....

   ....

   ....

   ....

   ....

2) There are 4 (A,B,C, and D) sections in class X of a secondary school. You are asked to find the average income of the parents of the students of section A and C. Which sampling method will be used from the following?

   a) Simple random sampling; b) Systematic sampling; c) Stratified sampling; d) Cluster sampling.

   ....

3) The employees of a company are classified into 4 groups (A,B,C and D) on the basis of their salary structure. You are asked to find the average salary income of the employees working in the company. What is the sampling method to be used?

   a) Simple random sampling;   b) Systematic sampling;   c) Stratified sampling;   d) Quota sampling.

   ....

4) State true or false.

    a) A systematic sampling can be used even if all the units of the population are not available.

    b) A budget has been announced by the government. A TV journalist recorded the views of the people residing near his house. The sampling method that the TV journalist used is quota sampling.

....

4.6   SAMPLE SIZE

The question of how large a sample should be is a difficult one. Sample size can be determined by various factors (like time, funds, manpower, population size, purpose of study etc. For example, if the available funds for study are limited then the researcher may not be able to spend more than a fixed proportion of the total fund available with him/her. In general, sample size depends on the nature of the analysis to be performed, the desired precision of the estimates one wishes to achieve, number of variables that have to be examined simultaneously and how heterogeneous is the population spread. Moreover, technical considerations suggest that the required sample size is a function of the precision of the estimates one wishes to achieve, the variance of the population and statistical level of confidence one wishes to use. The higher the precision and confidence level required, the larger the sample size should be. Typical confidence levels are 95% and 99%, while a typical precision (significance) value is 1% or 5%. You will learn more about the confidence and precision levels in Unit 16 and Unit 17 of this course.

Once the researcher determines the desired degree of precision and confidence level, there are several formulas he/she can use to determine the sample size and interpretation of results depending on the plan of the study. Here we will discuss three of them.

1) If the researcher wishes to report the results as proportions of the sample responses, use the following formula.

$$n = \frac{P(1-P)}{\dfrac{A^2}{Z^2} + \dfrac{P(1-P)}{N}}$$

Where, n = Sample size.

   P = Estimated percentage of the population possessing attribute of interest.

   A = Accuracy desired, usually expressed as a decimal (i.e. 0.01, 0.05, etc.)

   Z = Standardization value indicating a confidence level (Z=1.96 at 95% confidence level and Z = 2.56 at 99% confidence level. See Unit 16 for more details.

   N = Population size (known or estimated)

2) If the researcher wishes to report the results as means of the sample responses, use the following formula.

$$n = \frac{\sigma^2}{\dfrac{A^2}{Z^2} + \dfrac{\sigma^2}{N}}$$

Where, n = Sample size.

       P = Estimated percentage of the population possessing attribute of interest.

       A = Accuracy desired, usually expressed as a decimal (i.e. 0.01, 0.05, etc.)

       Z = Standardization value indicating a confidence level (Z=1.96 at 95% confidence level and Z = 2.56 at 99% confidence level. See Unit 16 for more details.

       N = Population size (known or estimated)

3) If the researcher plans the results in a variety of ways or if he/she has difficulty in estimating the proportion or standard deviation of the attribute of interest, the following formula may be more useful.

$$n = \frac{NZ^2 \times .25}{[d^2 \times (N-1)] + [Z^2 \times .25]}$$

Where, n = Sample size required

       d = Accuracy precision level (i.e. 0.01, 0.05, 0.10 etc.)

       Z = Standardization value indicating a confidence level (Z = 1.96 at 95% confidence level and Z = 2.56 at 99% confidence level. See Unit 16 for more details.

       N = Population size (known or estimated).

For example, if the population size (N) is 1000 and you wish a 95% confidence level and ±5% precision level (d=0.05 and Z=1.96) then the sample size (n):

$$n = \frac{1000 \times 1.96^2 \times 0.25}{(0.05^2 \times 999) + (1.96^2 \times 0.25)} = 277.7 \text{ or say } 280$$

4.7 SAMPLING AND NON-SAMPLING ERRORS

The quality of a research project depends on the accuracy of the data collected and its representation to the population. There are two broad sources of errors. These are sampling errors and non-sampling errors.

4.7.1 Sampling Errors

The principal sources of sampling errors are the sampling method applied, and the sample size. This is due to the fact that only a part of the population is covered in the sample. The magnitude of the sampling error varies from one sampling method to the other, even for the same sample size. For example, the sampling error associated with simple random sampling will be greater than stratified random sampling if the population is heterogeneous in nature.

Intuitively, we know that the larger the sample the more accurate the research. In fact, the sampling error varies with samples of different sizes. Increasing the sample size decreases the sampling error.

The following Figure gives an approximate relationship between sample size and sampling error. Study the following figure carefully.
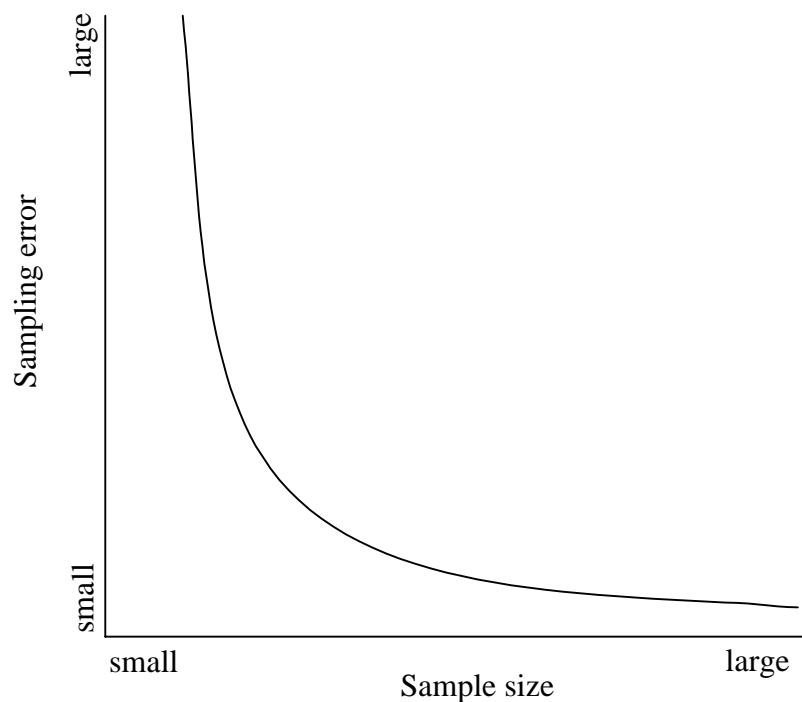

Fig.: 4.1

4.7.2 Non-Sampling Errors

The non-sampling errors arise from faulty research design and mistakes in executing research. There are many sources of non-sampling errors which may be broadly classified as: (a) respondent errors, and (b) administrative errors.

a) **Respondent Errors**: If the respondents co-operate and give the correct information the objectives of the researcher can be easily accomplished. However, in practice, this may not happen. The respondents may either refuse to provide information or even if he/she provides information it may be biased.

If the respondent fails to provide information, we call it as ***non-response error***. Although this problem is present in all types of surveys, the problem is more acute in mailed surveys. Non-response also leads to some extreme situations like those respondents who are willing to provide information are over represented while those who are indifferent are under-represented in the sample. In order to minimise the non-response error the researcher often seeks to re-contact with the non-respondents if they were not available earlier.

If the researcher finds that the non-response rate is more in a particular group of respondents (for example, higher income groups) additional efforts should be made to obtain data from these under-represented groups of the population. For example, for these people who are not responding to the mailed questionnaires, personal interviews may be conducted to obtain data. In a mailed questionnaire the researcher never knows whether the respondent really refused to provide data or was simply indifferent. There are several techniques which help to encourage respondents to reply. You must have already learned these techniques in Unit 3 of this course.

Response bias occurs when the respondent may not give the correct information and try to mislead the investigator in a certain direction. The

respondents may consciously or unconsciously misrepresent the truth. For example, if the investigator asks a question on the income of the respondent he/she may not give the correct information for obvious reasons. Or the investigator may not be able to put a question that is sensitive (thus avoiding embarrassment). This may arise from the problems in designing the questionaire and the content of questions. Respondents who must understand the questions may unconsciously provide biased information.

The response bias may also occur because the interviewer's presence influences respondents to give untrue or modified answers. The respondents/interviewers tendency is to please the other person rather than provide/elicit the correct information.

b) **Administrative Errors:** The errors that have arisen due to improper administration of the research process are called administrative errors. There are four types of administrative errors. These are as follows:

  i)   sample selection error,
  ii) investigator error,
  iii) investigator cheating, and
  iv. data processing error.

  i) **Sample Selection Error:** It is difficult to execute a sampling plan. For example, we may plan to use systematic sampling plan in a market research study of a new product and decide to interview every 5th customer coming out of a consumer store. If the day of interview happened to be a working day then we are excluding all those consumers who are working. This may lead to an error because of the unrepresentative sample selection.

  ii) **Investigator Error:** When the investigator interviews the respondent, he/she may fail to record the information correctly or may fail to cross check the information provided by the respondent. Therefore, the error may arise due to the way the investigator records the information.

  iii) **Investigator Cheating:** Some times the investigator may try to fake the data even without meeting the concerned respondents. There should be some mechanism to crosscheck this type of faking by the investigator.

  iv) **Data Processing Error:** Once the data is collected the next job the researcher does is edit, code and enter the data into a computer for further processing and analysis. The errors can be minimised by careful editing, coding and entering the data into a computer.

4.7.3   Control of Errors

In the above two sections we have identified the most significant sources of errors. It is not possible to eliminate completely the sources of errors. However, the researcher's objective and effort should be to minimise these sources of errors as much as possible. There are ways of reducing the errors. Some of these are:

(a) designing and executing a good questionnaire; (b) selection of appropriate sampling method; (c) adequate sample size; (d) employing trained investigators to collect the data; and (e) care in editing, coding and entering the data into the computer. You have already learned the above ways of controlling the errors in Unit 3 and in this Unit.

1) The size of a population is 10000. You wish to have a 99% confidence level and ±5% precision level. What is the sample size required?

   ....

   ....

   ....

2) As the sample size increases, the sampling error:

   a) Increases   b) Decreases   c) Remains constant

   ....

3) The sampling errors arise due to:

   a) The investigator's bias   b) The data processing problem

   c) The respondent's bias   d) The sampling method applied

   ....

4.8   LET US SUM UP

A sample is a subset of population whose characteristics are studied to know the information about the population. A complete survey of population is called census. When compared with census, sampling is less expensive, requires less time and other resources and is more accurate when samples are taken properly. Also, sampling is the only alternative when the measurement of population units is destructive in nature.

There are two broad categories of sampling methods. These are: (a) random sampling methods, and (b) non-random sampling methods. The random sampling methods are based on the chance of including the units of population in a sample.

Some of the sampling methods covered in this Unit are: (a) simple random sampling, (b) systematic random sampling, (c) stratified random sampling, (d) cluster sampling, and (e) multistage sampling. With an appropriate sampling plan and selection of random sampling method the sampling error can be minimised. The non-random sampling methods include: (a) convenience sampling, (b) judgment sampling, and (c) Quota sampling. These methods may be convenient to the researcher to apply. These methods may not provide a representative sample to the population and there are no scientific ways to check the sampling errors.

There are two major sources of errors in survey research. These are: (a) sampling errors, and (b) non-sampling errors. The sampling errors arise because of the fact that the sample may not be a representative sample to the population. Two major sources of non-sampling errors are due to: (a) non-response on the part of respondent and/or respondent's bias in providing correct information, and (b) administrative errors like design and implementation of questionnaire, investigators' bias, and data processing errors.

It may not be possible to completely eliminate the sampling and non-sampling errors. However, there are some ways to minimise these errors. These are:

(a) designing a good questionnaire, (b) selection of appropriate sampling method, (c) adequate sample size, (d) employing trained investigators and, (e) care in data processing.

4.9   KEY WORDS

Administrative Errors : The administrative errors arise due to improper administration of the research.

Census : A complete survey of population is called census.

Convenient Sampling : Here the units of the population are included in the sample as per the convenience of the researcher.

Cluster Sampling: In cluster sampling method we divide the population into groups called clusters, selective sample of clusters using simple random sampling and then cover all the units in each of the clusters included in the sample.

Judgment Sampling: In this sampling method the selection of sample is based on the researcher's judgment about some appropriate characteristics required of the sample units.

Multi-stage Sampling: Here we select the sample units in a number of stages using one or more random sampling methods.

Non-sampling Errors : The non-sampling errors arise from faulty research design and mistakes in executing the research.

Non-random Sampling/Non-Probability Sampling : In this sampling method the probability of any particular unit of the population being included in the sample is unknown.

Parameters : The numerical characteristics of a population are called parameters.

Quota Sampling : In this sampling method the samples are selected on the basis of some parameters such as age, gender, geographical region, education, income, occupation etc.

Random Sampling/Probability Sampling : If all the units of the population have a chance of being chosen in the sample, the sampling method is called random sampling/probability sampling.

Respondent Errors : The respondent errors arise due to failure of the respondent to provide correct information.

Sample : A sample is a representative set of population.

Sampling Errors : The sampling errors arise because we cover only a part of the population.

Simple Random Sampling : This is one of the basic methods of random sampling where each unit in the population has equal chance of being included in the sample.

Stratified Sampling : The stratified sampling method is used when the population is heterogeneous. Here the population is divided into some homogeneous groups called stratums.

Systematic Sampling : In systematic sampling the sample units are selected from the population at equal intervals in terms of time, space or order.

4.10 ANSWERS TO SELF ASSESSMENT EXERCISES

A. 2) a) sample survey; b) census; c) sample survey.

   3) a) biased; b) cost; c) sufficiently, economical.

B. 1) i) Selected sample using simple random sampling

283, 581, 359, 538, 435, 103, 428, 042, 678, 868,

699, 163, 652, 845, 308, 859, 292, 587, 960, 159,

370, 065, 580, 285, 884, 410, 135, 495, 032, 242

   ii) 39**971**, 49**978**, 84**904**, 87**984**, 55**960**, 83**928**

The population size is 900 and these random numbers fall outside the population range of 000 to 899.

   2) Cluster sampling

   3) Stratified sampling

   4) a) true

      b) false, it is convenience sampling

C. 1) The required sample size is 370

   2) Decreases

   3) Sampling method applied

4.11 TERMINAL QUESTIONS

1) What is the difference between random sampling and non-random sampling?

2) List some of the situations where (a) sampling is more appropriate than census and (b) census is more appropriate than sampling.

3) What are the advantages and disadvantages of stratified random sampling?

4) What are the ways to control survey errors?

5) What are the advantages of sampling over census?

6) Discuss the method of cluster sampling. What is the difference between cluster sampling and stratified random sampling?

7) The total population is 5000 and you wish a 99% confidence level and a ±5% precision level. What is the sample size required?

8) A certain population is divided into 4 stratums so that $N_1 = 4000$, $N_2 = 6000$, $N_3 = 7000$, $N_4 = 3000$. The respective stratum standard deviations are $\sigma_1 = 2.0$, $\sigma_2 = 4.0$, $\sigma_3 = 3.0$, $\sigma_4 = 6.0$. How should a sample size of 300 be allocated to four stratums using: (a) proportional and (b) disproportional methods.

9) Discuss the sources of sampling and non-sampling errors.

10) What are the essentials of a good sample?

> **Note:** These questions/exercises will help you to understand the unit better. Try to write answers for them. But do not submit your answers to the university for assessment. These are for your practice only.

4.12   FURTHER READING

The following text books may be used for more indepth study on the topics dealt with in this unit.

Gupta, C.B., & Vijay Gupta, *An Introduction to Statistical Methods,* Vikas Publishing House Pvt. Ltd., New Delhi.

Kothari, C.R.(2004) *Research Methodology Methods and Techniques*, New Age International (P) Ltd., New Delhi.

Levin, R.I. and D.S. Rubin. (1999) *Statistics for Management*, Prentice-Hall of India, New Delhi

Mustafi, C.K.(1981) *Statistical Methods in Managerial Decisions*, Macmillan, New Delhi

UNIT 5 MEASUREMENT AND SCALING TECHNIQUES

5.0 OBJECTIVES

After studying this unit, you should be able to:

l explain the concepts of measurement and scaling,

l discuss four levels of measurement scales,

l classify and discuss different scaling techniques, and

l select an appropriate attitude measurement scale for your research problem.

5.1 INTRODUCTION

As we discussed earlier, the data consists of quantitative variables like price, income, sales etc., and qualitative variables like knowledge, performance, character etc. The qualitative information must be converted into numerical form for further analysis. This is possible through measurement and scaling techniques. A common feature of survey based research is to have respondent's feelings, attitudes, opinions, etc. in some measurable form. For example, a bank manager may be interested in knowing the opinion of the customers about the services provided by the bank. Similarly, a fast food company having a network in a city may be interested in assessing the quality and service provided by them. As a researcher you may be interested in knowing the attitude of the people towards the government announcement of a metro rail in Delhi. In this unit we will discuss the issues related to measurement, different levels of measurement scales, various types of scaling techniques and also selection of an appropriate scaling technique.

5.2 MEASUREMENT AND SCALING

Before we proceed further it will be worthwhile to understand the following two terms: (a) Measurement, and (b) Scaling.

a) **Measurement:** Measurement is the process of observing and recording the observations that are collected as part of research. The recording of the

observations may be in terms of numbers or other symbols to characteristics of objects according to certain prescribed rules. The respondent's, characteristics are feelings, attitudes, opinions etc. For example, you may assign '1' for Male and '2' for Female respondents. In response to a question on whether he/she is using the ATM provided by a particular bank branch, the respondent may say 'yes' or 'no'. You may wish to assign the number '1' for the response yes and '2' for the response no. We assign numbers to these characteristics for two reasons. First, the numbers facilitate further statistical analysis of data obtained. Second, numbers facilitate the communication of measurement rules and results. The most important aspect of measurement is the specification of rules for assigning numbers to characteristics. The rules for assigning numbers should be standardised and applied uniformly. This must not change over time or objects.

b) **Scaling:** Scaling is the assignment of objects to numbers or semantics according to a rule. In scaling, the objects are text statements, usually statements of attitude, opinion, or feeling. For example, consider a scale locating customers of a bank according to the characteristic "agreement to the satisfactory quality of service provided by the branch". Each customer interviewed may respond with a semantic like 'strongly agree', or 'somewhat agree', or 'somewhat disagree', or 'strongly disagree'. We may even assign each of the responses a number. For example, we may assign strongly agree as '1', agree as '2' disagree as '3', and strongly disagree as '4'. Therefore, each of the respondents may assign 1, 2, 3 or 4.

5.3 ISSUES IN ATTITUDE MEASUREMENT

When a researcher is interested in measuring the attitudes, feelings or opinions of respondents he/she should be clear about the following:

a) What is to be measured?

b) Who is to be measured?

c) The choices available in data collection techniques

The first issue that the researcher must consider is 'what is to be measured'? The definition of the problem, based on our judgments or prior research indicates the concept to be investigated. For example, we may be interested in measuring the performance of a fast food company. We may require a precise definition of the concept on how it will be measured. Also, there may be more than one way that we can measure a particular concept. For example, in measuring the performance of a fast food company we may use a number of measures to indicate the performance of the company. We may use sales volume in terms of value of sales or number of customers or spread of network of the company as measures of performance. Further, the measurement of concepts requires assigning numbers to the attitudes, feelings or opinions. The key question here is that on what basis do we assign the numbers to the concept. For example, the task is to measure the agreement of customers of a fast food company on the opinion of whether the food served by the company is tasty, we create five categories: (1) strongly agree, (2) agree, (3) undecided, (4) disagree, (5) strongly disagree. Then we may measure the response of respondents. Suppose if a respondent states 'disagree' with the statement that 'the food is tasty', the measurement is 4.

The second important issue in measurement is that, who is to be measured? That means who are the people we are interested in. The characteristics of the people such as age, sex, education, income, location, profession, etc. may

have a bearing on the choice of measurement. The measurement procedure must be designed keeping in mind the characteristics of the respondents under consideration.

The third issue in measurement is the choice of the data collection techniques. In Unit 3, you have already learnt various methods of data collection.Normally, questionnaires are used for measuring attitudes, opinions or feelings.

5.4 LEVELS OF MEASUREMENT SCALES

The level of measurement refers to the relationship among the values that are assigned to the attributes, feelings or opinions for a variable. For example, the variable 'whether the taste of fast food is good' has a number of attributes, namely, very good, good, neither good nor bad, bad and very bad. For the purpose of analysing the results of this variable, we may assign the values 1, 2, 3, 4 and 5 to the five attributes respectively. The level of measurement describes the relationship among these five values. Here, we are simply using the numbers as shorter placeholders for the lengthier text terms. We don't mean that higher values mean 'more' of something or lower values mean 'less' of something. We don't assume that 'good' which has a value of 2 is twice of 'very good' which has a value of 1. We don't even assume that 'very good' which is assigned the value '1' has more preference than 'good' which is assigned the value '2'. We simply use the values as a shorter name for the attributes, opinions, or feelings. The assigned values
of attributes allow the researcher more scope for further processing of data and statistical analysis.

Typically, there are four levels of measurement scales or methods of assigning numbers: (a) Nominal scale, (b) Ordinal scale, (c) Interval scale, and (d) Ratio scale.

a) **Nominal Scale** is the crudest among all measurement scales but it is also the simplest scale. In this scale the different scores on a measurement simply indicate different categories. The nominal scale does not express any values or relationships between variables. For example, labelling men as '1' and women as '2' which is the most common way of labelling gender for data recording purpose does not mean women are 'twice something or other' than men. Nor it suggests that men are somehow 'better' than women. Another example of nominal scale is to classify the respondent's income into three groups: the highest income as group 1. The middle income as group 2, and the low-income as group 3. The nominal scale is often referred to as a categorical scale. The assigned numbers have no arithmetic properties and act only as labels. The only statistical operation that can be performed on nominal scales is a frequency count. We cannot determine an average except mode.

In designing and developing a questionnaire, it is important that the response categories must include all possible responses. In order to have an exhaustive number of responses, you might have to include a category such as 'others', 'uncertain', 'don't know', or 'can't remember' so that the respondents will not distort their information by forcing their responses in one of the categories provided. Also, you should be careful and be sure that the categories provided are mutually exclusive so that they do not overlap or get duplicated in any way.

b) **Ordinal Scale** involves the ranking of items along the continuum of the characteristic being scaled. In this scale, the items are classified according to

whether they have more or less of a characteristic. For example, you may wish to ask the TV viewers to rank the TV channels according to their preference and the responses may look like this as given below:

TV Channel	Viewers preferences
Doordarshan-1	1
Star plus	2
NDTV News	3
Aaaj Tak TV	4

The main characteristic of the ordinal scale is that the categories have a logical or ordered relationship. This type of scale permits the measurement of degrees of difference, (that is, 'more' or 'less') but not the specific amount of differences (that is, how much 'more' or 'less'). This scale is very common in marketing, satisfaction and attitudinal research.

Another example is that a fast food home delivery shop may wish to ask its customers:

How would you rate the service of our staff?

(1) Excellent • (2) Very Good • (3) Good • (4) Poor • (5) Worst •

Suppose respondent X gave the response 'Excellent' and respondent Y gave the response 'Good', we may say that respondent X thought that the service provided better than respondent Y to be thought. But we don't know how much better and even we can't say that both respondents have the same understanding of what constitutes 'good service'.

In marketing research, ordinal scales are used to measure relative attitudes, opinions, and preferences. Here we rank the attitudes, opinions and preferences from best to worst or from worst to best. However, the amount of difference between the ranks cannot be found out. Using ordinal scale data, we can perform statistical analysis like Median and Mode, but not the Mean.

c) **Interval Scale** is a scale in which the numbers are used to rank attributes such that numerically equal distances on the scale represent equal distance in the characteristic being measured. An interval scale contains all the information of an ordinal scale, but it also one allows to compare the difference/distance between attributes. For example, the difference between '1' and '2' is equal to the difference between '3' and '4'. Further, the difference between '2' and '4' is twice the difference between '1' and '2'. However, in an interval scale, the zero point is arbitrary and is not true zero. This, of course, has implications for the type of data manipulation and analysis. We can carry out on data collected in this form. It is possible to add or subtract a constant to all of the scale values without affecting the form of the scale but one cannot multiply or divide the values. Measuring temperature is an example of interval scale. We cannot say 40^0C is twice as hot as 20^0C. The reason for this is that 0^0C does not mean that there is no temperature, but a relative point on the Centigrade Scale. Due to lack of an absolute zero point, the interval scale does not allow the conclusion that 40^0C is twice as hot as 20^0C.

Interval scales may be either in numeric or semantic formats. The following are two more examples of interval scales one in numeric format and another in semantic format.

i) **Example of Interval Scale in Numeric Format**

Food supplied is:						Indicate your score on
Fresh	1	2	3	4	5	the concerned blank and circle the appropriate number on each line.
Tastes good	1	2	3	4	5	
Value for money	1	2	3	4	5	
Attractive packaging	1	2	3	4	5	
Prompt time delivery	1	2	3	4	5	

ii) **Example of Interval Scale in Semantic Format**

Please indicate your views on the food supplied by XXX Fast Food Shop by scoring them on a five points scale from 1 to 5 (that is, 1=Excellent, 2=Very Good, 3=Good, 4=Poor, 5=Worst). Indicate your views by ticking the appropriate responses below:

Food supplied is:	Excellent	Very Good	Good	Poor	Worst
Fresh					
Tastes good					
Value for money					
Attractive packaging					
Prompt time delivery					

The interval scales allow the calculation of averages like Mean, Median and Mode and dispersion like Range and Standard Deviation.

d) **Ratio Scale** is the highest level of measurement scales. This has the properties of an interval scale together with a fixed (absolute) zero point. The absolute zero point allows us to construct a meaningful ratio. Examples of ratio scales include weights, lengths and times. In the marketing research, most counts are ratio scales. For example, the number of customers of a bank's ATM in the last three months is a ratio scale. This is because you can compare this with previous three months. Ratio scales permit the researcher to compare both differences in scores and relative magnitude of scores. For example, the difference between 10 and 15 minutes is the same as the difference between 25 and 30 minutes and 30 minutes is twice as long as 15 minutes. Most financial research that deals with rupee values utilizes ratio scales. However, for most behavioural research, interval scales are typically the highest form of measurement. Most statistical data analysis procedures do not distinguish between the interval and ratio properties of the measurement scales and it is sufficient to say that all the statistical operations that can be performed on interval scale can also be performed on ratio scales.

Now you must be wondering why you should know the level of measurement. Knowing the level of measurement helps you to decide on how to interpret the data. For example, when you know that a measure is nominal then you know that the numerical values are just short codes for longer textual names. Also, knowing the level of measurement helps you to decide what statistical analysis is appropriate on the values that were assigned. For example, if you know that a measure is nominal, then you would not need to find mean of the data values or perform a t-test on the data. (t-test will be discussed in Unit-16 in the course).

It is important to recognise that there is a hierarchy implied in the levels of measurement. At lower levels of measurement, assumptions tend to be less restrictive and data analyses tend to be less sensitive. At each level up the hierarchy, the current level includes all the qualities of the one below it and adds something new. In general, it is desirable to have a higher level of measurement (that is, interval or ratio) rather than a lower one (that is, nominal or ordinal).

Self Assessment Exercise-A

1) The main difference between interval scale and the ratio scale in terms of their properties is:

   ....

   ....

   ....

2) Why should the researcher know the level of measurement?

   ....

   ....

   ....

3) What are the main statistical limitations of nominal scale?

   ....

   ....

   ....

4) Indicate whether the following measures are nominal, ordinal, interval or ratio scales?

   a) Social status of a respondent: ....

   b) Stock market prices: ....

   c) The ranks obtained by students: ....

   d) The Fahrenheit scale for measuring temperature: ....

5.5 TYPES OF SCALING TECHNIQUES

The various types of scaling techniques used in research can be classified into two categories: (a) comparative scales, and (b) Non-comparative scales. In **comparative scaling**, the respondent is asked to compare one object with another. For example, the researcher can ask the respondents whether they prefer brand A or brand B of a detergent. On the other hand, in non-comparative scaling respondents need only evaluate a single object. Their evaluation is independent of the other object which the researcher is studying. Respondents using a non-comparative scale employ whatever rating standard seems appropriate to them. Non-comparative techniques consist of continuous and itemized rating scales. Figure 5.1 shows the classification of these scaling techniques.

Figure 5.1: Scaling Techniques
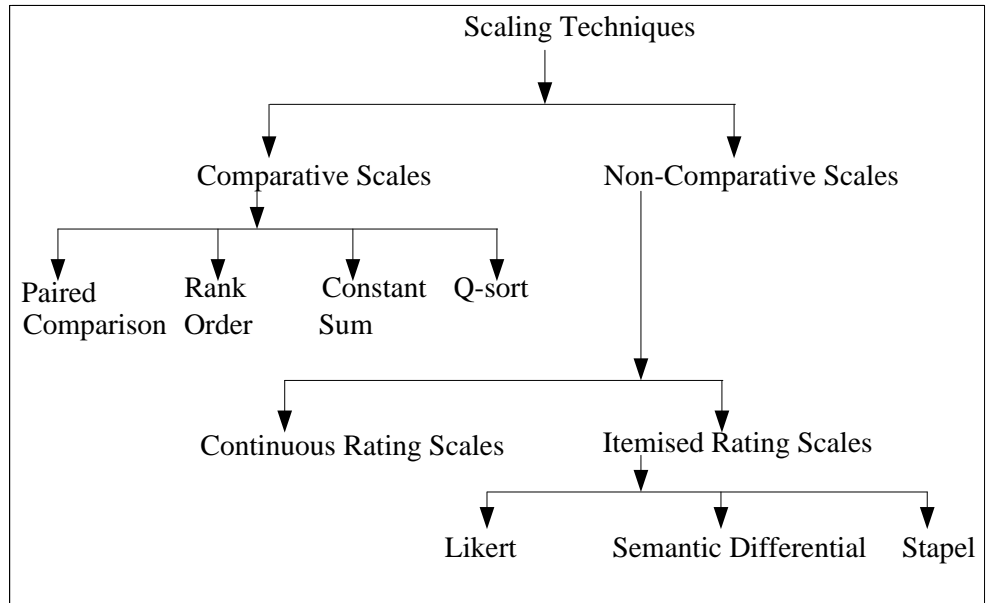
5.5.1 Comparative Scales

The comparative scales can further be divided into the following four types of scaling techniques: (a) Paired Comparison Scale, (b) Rank Order Scale, (c) Constant Sum Scale, and (d) Q-sort Scale.

a) **Paired Comparison Scale:** This is a comparative scaling technique in which a respondent is presented with two objects at a time and asked to select one object (rate between two objects at a time) according to some criterion. The data obtained are ordinal in nature. For example, there are four types of cold drinks - Coke, Pepsi, Sprite, and Limca. The respondents can prefer Pepsi to Coke or Coke to Sprite, etc. In all we can have the following six comparisons.

Coke–Pepsi

Coke–Sprite

Coke–Limca

Pepsi–Sprite

Pepsi–Limca

Sprite–Limca

In general, with n brands we have $\dfrac{n(n-1)}{2}$ paired comparisons. The following is the data recording format using the paired comparisons.

Table 5.1

Brand	Coke	Pepsi	Sprite	Limca
Coke	—	√		
Pepsi		—		
Sprite	√	√	—	
Limca	√	√	√	—
No. of times preferred	2	3	1	0

A √ in a particular box means that the brand in that column was preferred over the brand in the corresponding row. In the above recording, Coke was preferred over Sprite, Coke over Limca, in this case the number of times coke preferred was 2 times. Similarly, Pepsi over Coke, Pepsi over Sprite, Pepsi over Limca, in this case Pepsi was 3 time preferred. Thus, the number of times a brand was preferred is obtained by summing the √ s in each column.

The following table gives paired comparison of data (assumed) for four brands of cold drinks.

Table 5.2

Brand	Coke	Pepsi	Sprite	Limca
Coke	–	0.90	0.64	0.14
Pepsi	0.10	–	0.32	0.02
Sprite	0.36	0.68	–	0.15
Limca	0.86	0.98	0.85	–

The entries in the boxes represent the proportion of respondents preferring 'column brand' and to 'row' brand. For example, 90% prefer Pepsi to Coke and only 10% prefer Coke to Pepsi, etc.

Paired comparison is useful when the number of brands are limited, since it requires direct comparison and overt choice. One of the disadvantages of paired comparison scale is violation of the assumption of transitivity may occur. For example, in our example (Table 5.1) the respondent preferred Coke 2 times, Pepsi 3 times, Sprite 1 time, and Limca 0 times. That means, preference-wise, Pepsi >Coke, Coke >Sprite, and Sprite >Limca. However, the number of times Sprite was preferred should not be that of Coke. In other words, if A>B and B >C then C >A should not be possible. Also, the order in which the objects are presented may bias the results. The number of items/brands for comparison should not be too many. As the number of items increases, the number of comparisons increases geometrically. If the number of comparisons is too large, the respondents may become fatigued and no longer be able to carefully discriminate among them. The other limitation of paired comparison is that this scale has little resemblance to the market situation, which involves selection from multiple alternatives. Also, respondents may prefer one item over certain others, but they may not like it in an absolute sense.

b) **Rank Order Scale:** This is another type of comparative scaling technique in which respondents are presented with several items simultaneously and asked to rank them in the order of priority. This is an ordinal scale that describes the favoured and unfavoured objects, but does not reveal the distance between the objects. For example, if you are interested in ranking the preference of some selected brands of cold drinks, you may use the following format for recording the responses.

Table 5.3: Preference of cold drink brands using rank order scaling

Instructions: Rank the following brands of cold drinks in order of preference. Begin by picking out the one brand you like most and assign it a number 1. Then find the second most preferred brand and assign it a number 2. Continue this procedure until you have ranked all the brands of cold drinks in order of preference. The least preferred brand should be assigned a rank of 4. Also remember no two brands receive the same rank order.

Format:

Brand	Rank
(a) Coke	3
(b) Pepsi	1
(c) Limca	2
(d) Sprite	4

Like paired comparison, the rank order scale, is also comparative in nature. The resultant data in rank order is ordinal data. This method is more realistic in obtaining the responses and it yields better results when direct comparison are required between the given objects. The major disadvantage of this technique is that only ordinal data can be generated.

c) **Constant Sum Scale:** In this scale, the respondents are asked to allocate a constant sum of units such as points, rupees, or chips among a set of stimulus objects with respect to some criterion. For example, you may wish to determine how important the attributes of price, fragrance, packaging, cleaning power, and lather of a detergent are to consumers. Respondents might be asked to divide a constant sum to indicate the relative importance of the attributes using the following format.

Table 5.4: Importance of detergent attributes using a constant sum scale

Instructions: Between attributes of detergent please allocate 100 points among the attributes so that your allocation reflects the relative importance you attach to each attribute. The more points an attribute receives, the more important the attribute is. If an attribute is not at all important, assign it zero points. If an attribute is twice as important as some other attribute, it should receive twice as many points.

Format:

Attribute	Number of Points
(a) Price	50
(b) Fragrance	05
(c) Packaging	10
(d) Cleaning Power	30
(e) Lather	05
Total Points	100

"If an attribute is assigned a higher number of points, it would indicate that the attribute is more important." From the above Table, the price of the detergent is

the most important attribute for the consumers followed by cleaning power, packaging. Fragrance and lather are the two attributes that the consumers cared about the least but preferred equally." The advantage of this technique is saving time. However, there are two main disadvantages. The respondents may allocate more or fewer points than those specified. The second problem is rounding off error if too few attributes are used and the use of a large number of attributes may be too taxing on the respondent and cause confusion and fatigue.

d) **Q-Sort Scale:** This is a comparative scale that uses a rank order procedure to sort objects based on similarity with respect to some criterion. The important characteristic of this methodology is that it is more important to make comparisons among different responses of a respondent than the responses between different respondents. Therefore, it is a comparative method of scaling rather than an absolute rating scale. In this method the respondent is given statements in a large number for describing the characteristics of a product or a large number of brands of a product. For example, you may wish to determine the preference from among a large number of magazines. The following format shown in Table 5.5 may be given to a respondent to obtain the preferences.

Table 5.5: Preference of Magazines Using Q-Sort Scale Procedure

Instructions: The bag given to you contain pictures of 90 magazines. Please choose 10 magazines you 'prefer most', 20 magazines you 'like', 30 magazines to which you are 'neutral (neither like nor dislike)', 20 magazines you 'dislike', and 10 magazines you 'prefer least'. Please list the sorted magazine names in the respective columns of the form provided to you.

Format:

Prefer Most	Like	Neutral	Dislike	Prefer Least
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____
(10)	_____	_____	_____	(10)
	_____	_____	_____	
	_____	_____	_____	
	_____	_____	_____	
	_____	_____	_____	
	_____	_____	_____	
	_____	_____	_____	
	_____	_____	_____	
	_____	_____	_____	
	_____	_____	_____	
	(20)	_____	(20)	
		_____		
		_____		
		_____		
		_____		
		_____		
		_____		
		_____		
		_____		
		_____		
		(30)		

Note that the number of responses to be sorted should not be less than 60 or not more than 140. A reasonable range is 60 to 90 responses that result in a normal or quasi-normal distribution. This method is faster and less tedious than paired comparison measures. It also forces the subject to conform to quotas at each point of scale so as to yield a quasi-normal distribution. The utility of Q-sort in marketing research is to derive clusters of individuals who display similar preferences, thus representing unique market segments.

5.5.2   Non-Comparative Scales

The non-comparative scaling techniques can be further divided into: (a) Continuous Rating Scale, and (b) Itemised Rating Scale.

a) **Continuous Rating Scales**

It is very simple and highly useful. In continuous rating scale, the respondent's rate the objects by placing a mark at the appropriate position on a continuous line that runs from one extreme of the criterion variable to the other. Examples of continuous rating scale are given below:

Question: How would you rate the TV advertisement as a guide for buying?

Scale Type A

Strongly agree ———————————————————————— Strongly disagree

Scale Type B

Strongly disagree ———————————————————————— Strongly agree

Scale Type C

Strongly agree | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | Strongly disagree

Scale Type D

Strongly disagree | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Strongly agree
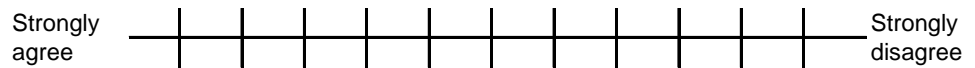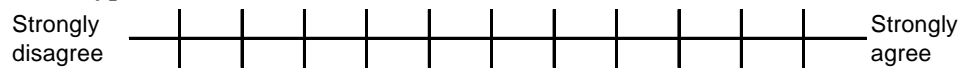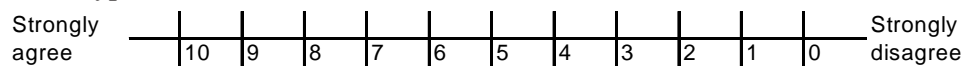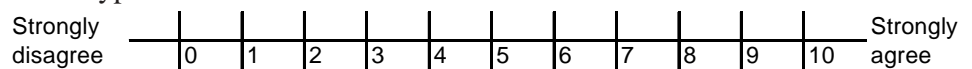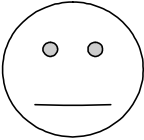
When scale type A and B are used, the respondents score is determined either by dividing the line into as many categories as desired and assigning the respondent a score based on the category into which his/her mark falls, or by measuring distance, in millimeters, centimeters, or inches from either end of the scale. Which ever of the above continuous scale is used, the results are normally analysed as interval scaled.

b) **Itemised Rating Scales**

Itemised rating scale is a scale having numbers or brief descriptions associated with each category. The categories are ordered in terms of scale position and the respondents are required to select one of the limited number of categories that best describes the product, brand, company, or product attribute being rated. Itemised rating scales are widely used in marketing research.

The itemised rating scales can be in the form of : (a) graphic, (b) verbal, or (c) numeric as shown below:

Itemised Graphic Scale	Itemised Verbal Scale	Itemised Numeric Scale
Favourable	Completely satisfied	−5
		−4
	Somewhat satisfied	−3
		−2
		−1
Indifferent	Neither satisfied nor dissatisifed	0
		+1
	Somewhat dissatisfied	+2
		+3
Unfavourable	Completely dissatisfied	+4
		+5

Some rating scales may have only two response categories such as : agree and disagree. Inclusion of more response categories provides the respondent more flexibility in the rating task. Consider the following questions:

1. *How often do you visit the supermarket located in your area of residence?*

   • Never,   • Rarely,   • Sometimes,   • Often,   • Very often

2. In your case how important is the price of brand X shoes when you buy them?

   • Very important,   • Fairly important,   • Neutral,   • Not so important

Each of the above category scales is a more sensitive measure than a scale with only two responses since they provide more information.

Wording is an extremely important factor in the usefulness of itemised scales. Table 5.6 shows some common wordings for categories used in itemised scales.

Table 5.6: Some common words for categories used in Itemised Rating scales

Quality:				
Excellent	Good	Not decided	Poor	Worst
Very Good	Good	Neither good nor bad	Fair	Poor

Importance:				
Very Important	Fairly important	Neutral	Not so important	Not at all important

Interest:				
Very interested	Somewhat interested	Neither interested nor disinterested	Somewhat uninterested	Not very interested

Satisfaction:				
Completely satisfied	Somewhat satisfied	Neither satisfied nor dissatisfied	Somewhat dissatisfied	Completely dissatisfied

Frequency:				
All of the time	Very often	Often	Sometimes	Hardly ever
Very ofen	Often	Sometimes	Rarely	Never

Truth:				
Very true	Somewhat true	Not very true	Not at all true	

Purchase Interest:				
Definitely will buy	Probably will buy	Probably will not buy	Definitely will not buy	

Level of Agreement:				
Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree

Dependability:				
Completely dependable	Somewhat dependable	Not very dependable	Not at all dependable	

Style:				
Very stylish	Somewhat stylish	Not very stylish	Completely unstylish	

Cost:				
Extremely expensive	Expensive	Neither expensive nor inexpensive	Slightly inexpensive	Very inexpensive

Ease of use:				
Very ease to use	Somewhat easy to use	Not very easy to use	Difficult to use	

Modernity:				
Very modern	Somehwat modern	Neither modern nor old-fashioned	Somewhat old fashioned	Very old fashioned

Alert:				
Very alert	Alert	Not alert	Not at all alert	

In this section we will discuss three itemised rating scales, namely (a) Likert scale, (b) Semantic Differential Scale, and (c) Stapel Scale.

a) **Likert Scale:** In business research, the Likert scale, developed by Rensis Likert, is extremely popular for measuring attitudes, because, the method is simple to administer. With the Likert scale, the respondents indicate their own attitudes by checking how strongly they agree or disagree with carefully worded statements that range from very positive to very negative towards the attitudinal object. Respondents generally choose from five alternatives (say strongly agree, agree, neither agree nor disagree, disagree, strongly disagree).

Consider the following example of a study or measuring attitudes towards cricket.

	Strongly agree	Agree	Not sure	Disagree	Strongly disagree
It is more fun to play a tough, competitive cricket match than to play an easy one.	5	4	3	2	1

To measure the attitude, the researchers assign weights or scores to the alternative responses. In the above example the scores 5 to 1 are assigned to the responses. Strong agreement of the respondent indicates the most favourable attitudes on the statement, and the score 5 is assigned to it. On the other hand, strong disagreement of the respondent indicates the most unfavourable attitude on the statement, and the score 1 is assigned to it. If a negative statement towards the object is given, the corresponding scores would be reversed. In this case, the response 'strongly agree' will get a score of 1 and the response 'strongly disagree' will get a score of 5.

A Likert scale may include a number of items or statements. Each statement is assumed to represent an aspect of an attitudinal domain. For example, Table 5.7 shows the items in a Likert Scale to measure opinions on food products.

Table 5.7: A Likert Scale for studying opinions on food products.

	Strongly Agree	Agree	Neither Agree Nor Disagree	Disagree	Strongly Disagree
If the price of raw materials fall, firms too should reduce the price of the food products.	1	2	3	4	5
There should be uniform price through out the country for food products	1	2	3	4	5
The food companies should concentrate more on keeping hygiene while manufacturing food products.	1	2	3	4	5
The expiry dates should be printed on the food products before they are delivered to consumers in the market.	1	2	3	4	5
There should be government regulations on the firms in keeping acceptable quality and on the prices	1	2	3	4	5
Now-a-days most food companies are concerned only with profit making rather than taking care of quality.	1	2	3	4	5

Each respondent is asked to circle his opinion on a score against each statement. The final score for the respondent on the scale is the sum of their ratings for all the items. The very purpose of Likert's Scale is to ensure the final items evoke a wide response and discriminate among those with positive and negative attitudes. Items that are poor (because they lack clarity or elicit mixed response patterns) are detected from the final statement list. This will ensure us to discriminate between high positive scores and high negative scores. However, many business researchers do not follow this procedure and you may not be in a position to distinguish between high positive scores and high negative scores because all scores look alike. Hence a disadvantage of the Likert Scale is that it is difficult to know what a single summated score means. Many patterns of response to the various statements can produce the same total score. The other disadvantage of Likert Scale is that it takes longer time to complete than other itemised rating scales because respondents have to read each statement. Despite the above disadvantages, this scale has several advantages. It is easy to construct, administer and use.

b) **Semantic Differential Scale:** This is a seven point rating scale with end points associated with bipolar labels (such as good and bad, complex and simple) that have semantic meaning. The Semantic Differential scale is used for a variety of purposes. It can be used to find whether a respondent has a positive or negative attitude towards an object. It has been widely used in comparing brands, products and company images. It has also been used to develop advertising and promotion strategies and in a new product development study. Look at the following Table, for examples of Semantic Differential Scale.

Table 5.8: Examples of Semantic Differential Scale

Modern	—	—	—	—	—	—	—	Old-fashioned
Good	—	—	—	—	—	—	—	Bad
Clean	—	—	—	—	—	—	—	Dirty
Important	—	—	—	—	—	—	—	Unimportant
Expensive	—	—	—	—	—	—	—	Inexpensive
Useful	—	—	—	—	—	—	—	Useless
Strong	—	—	—	—	—	—	—	Weak
Quick	—	—	—	—	—	—	—	Slow

In the Semantic Differential scale only extremes have names. The extreme points represent the bipolar adjectives with the central category representing the neutral position. The in between categories have blank spaces. A weight is assigned to each position on the scale. The weights can be such as +3, +2, +1, 0, –1, –2, –3 or 7,6,5,4,3,2,1. The following is an example of Semantic Differential Scale to study the experience of using a particular brand of body lotion.

In my experience, the use of body lotion of Brand-X was:							
	+3	+2	+1	0	−1	−2	−3	
Useful	——	——	——	——	——	——	——	Useless
Attractive	——	——	——	——	——	——	——	Unattractive
Passive	——	——	——	——	——	——	——	Active
Beneficial	——	——	——	——	——	——	——	Harmful
Interesting	——	——	——	——	——	——	——	Boring
Dull	——	——	——	——	——	——	——	Sharp
Pleasant	——	——	——	——	——	——	——	Unpleasant
Cold	——	——	——	——	——	——	——	Hot
Good	——	——	——	——	——	——	——	Bad
Likable	——	——	——	——	——	——	——	Unlikable

In the semantic Differential scale, the phrases used to describe the object form a basis for attitude formation in the form of positive and negative phrases. The negative phrase is sometimes put on the left side of the scale and sometimes on the right side. This is done to prevent a respondent with a positive attitude from simply checking the left side and a respondent with a negative attitude checking on the right side without reading the description of the words.

The respondents are asked to check the individual cells depending on the attitude. Then one could arrive at the average scores for comparisons of different objects. The following Figure shows the experiences of 100 consumers on 3 brands of body lotion.
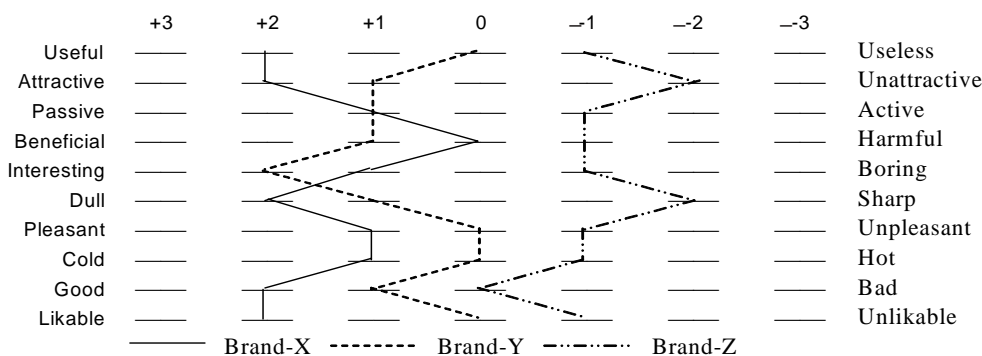
In the above example, first the individual respondent scores for each dimension are obtained and then the average scores of all 100 respondents, for each dimension and for each brand were plotted graphically. The maximum score possible for each brand is + 30 and the minimum score possible for each brand is –30. Brand-X has score +14. Brand-Y has score +7, and Brand-Z has score –11. From the scale we can identify which phrase needs improvement for each Brand. For example, Brand-X needs to be improved upon benefits and Brand-Y on pleasantness, coldness and likeability. Brand Z needs to be improved on all the attributes.

c) **Staple Scale:** The Stapel scale was originally developed to measure the direction and intensity of an attitude simultaneously. Modern versions of the Stapel scale place a single adjective as a substitute for the Semantic differential when it is difficult to create pairs of bipolar adjectives. The modified Stapel scale places a single adjective in the centre of an even number of numerical values (say, +3, +2, +1, 0, –1, –2, –3). This scale measures how close to or how distant from the adjective a given stimulus is perceived to be. The following is an example of a Staple scale.

Instructions: Select a plus number for words that you think describe personnel banking of a bank accurately. The more accurately you think the word describes the bank, the larger the plus number you should choose. Select a minus number for words you think do not describe the bank accurately. The less accurately you think the word describes the bank, the larger the minus number you should choose.

Format:

+5	+5
+4	+4
+3	+3
+2	+2
+1	+1
Friendly Personnel	Competitive Loan Rates
–1	–1
–2	–2
–3	–3
–4	–4
–5	–5

The following format shows an example of Stapel scale that illustrates respondents description on personnel banking of a bank.

	+4	+3	+2	+1	-1	-2	-3	-4
Fast Services	——	——	——	——	——	——	——	——
Friendly	——	——	——	——	——	——	——	——
Honest	——	——	——	——	——	——	——	——
Convenient Location	——	——	——	——	——	——	——	——
Convenient Hours	——	——	——	——	——	——	——	——
Dull	——	——	——	——	——	——	——	——
Good Services	——	——	——	——	——	——	——	——
High Saving Rates	——	——	——	——	——	——	——	——

Each respondent is asked to circle his opinion on a score against each phrase that describes the object. The final score of the respondent on a scale is the sum of their ratings for all the items. Also, the average score for each phrase is obtained by totaling the final score of all the respondents for that phrase divided by the number of respondents of the phrase. The following Figure shows the opinions of 100 respondents on two banks.

Bank-X —— Bank-Y ------
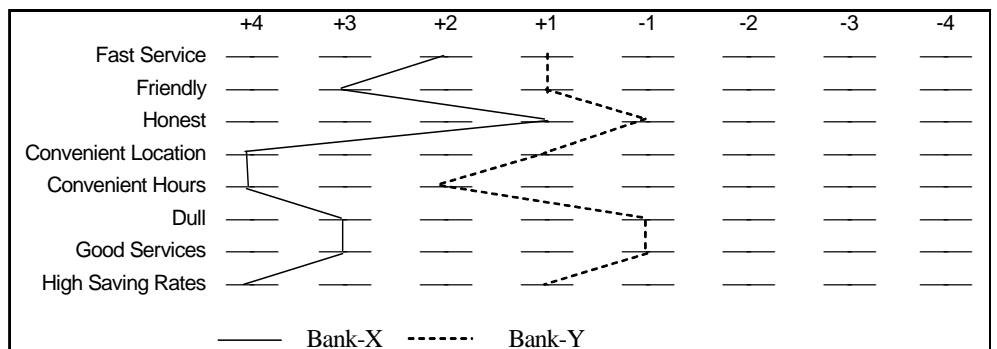
In the above example first the individual respondent's scores for each phrase that describes the selected bank are obtained and then the average scores of all

100 respondents for each phrase are plotted graphically. The maximum score possible for each bank is +32 and the minimum possible score for each brand is –32. In the example, Bank-X has score +24, and Bank-Y has score +3. From the scale we can identify which phrase needs improvement for each Bank.

The advantages and disadvantages of the Stapel scale are very similar to those for the Semantic differential scale. However, the Stapel scale tends to be easier to construct and administer, especially over telephone, since the Stapel scale does not call for the bipolar adjectives as does the Semantic differential scale. However, research on comparing the Stapel scale with Semantic differential scale suggests that the results of both the scales are largely the same.

5.6 SELECTION OF AN APPROPRIATE SCALING TECHNIQUE

In this unit, so far, you have learnt some of the important scaling techniques that are frequently used in attitudinal research for the measurement of attitudes. Each of these techniques has some advantages and disadvantages. Now you may ask which technique is more appropriate to use to measure attitudes. Virtually any technique can be used to measure the attitudes. But at the same time all techniques are not suitable for all purposes. As a general rule, you should use a scaling technique that will yield the highest level of information feasible in a given situation. Also, if possible the technique should permit you the use of a variety of statistical analysis. A number of issues decide the choice of scaling technique. Some significant issues are:

1) **Problem Definition and Statistical Analysis**: The Choice between ranking, sorting, or rating techniques is determined by the problem definition and the type of statistical analysis likely to be performed. For example, ranking provides only ordinal data that limits the use of statistical techniques.

2) **The Choice between Comparative and Non-comparative Scales**: Some times it is better to use a comparative scale rather than a non-comparative scale. Consider the following example:

How satisfied you are with the brand- X detergent that you are presently using?

Completely satisfied　*Somewhat satisfied*　*Neither satisfied nor dissatisifed*　*Somewhat dissatisfied*　*Completely dissatisfied*

This is a non-comparative scale since it deals with a single concept (the brand of a detergent). On the other hand, a comparative scale asks a respondent to rate a concept. For example, you may ask:

Which one of the following brands of detergent you prefer?

Brand-X　　Brand-Y

In this example you are comparing one brand of detergent with another brand. Therefore, in many situations, comparative scaling presents 'the ideal situation' as a reference for comparison with actual situation.

3) **Type of Category Labels:** We have discussed different types of category labels used in constructing measurement scales such as verbal categories and numeric categories. Many researchers use verbal categories since they believe that these categories are understood well by the respondents. The maturity and the education level of the respondents influences this decision.

4) **Number of Categories:** While there is no single, optimal number of categories, traditional guidelines suggest that there should be between five and nine categories. Also, if a neutral or indifferent scale response is possible for at least some of the respondents, an odd number of categories should be used. However, the researcher must determine the number of meaningful positions that are best suited for a specific problem.

5) **Balanced versus Unbalanced Scale:** In general, the scale should be balanced to obtain objective data.

6) **Forced versus Nonforced Categories:** In situations where the respondents are expected to have no opinion, the accuracy of data may be improved by a non forced scale that provides a 'no opinion' category.

Self Assessment Exercises B

1) In paired comparison, the order in which the objects are presented may _____ results.

2) A researcher wants to measure consumer preference between 7 brands of bath soap and has decided to use the Paired comparisons scaling technique. How many pairs of brands will the researcher present the respondents?: _____

3) In a semantic differential scale there are 20 scale items. Should all the positive adjectives be on the left side and all the negative adjectives be on the right side. Can you explain?

....

....

....

4) Indicate the type of scale used in the following examples.

   a) Do you favour or oppose the return of the ruling party in the next elections?

   (i) Favour   (ii) Neutral   (iii) Oppose

   ....

   b) Which one of the following pairs of companies would you like to invest your money?

   i) MTNL - Reliance

   ii) MTNL - BPL

   iii) Reliance - BPL

   c) Suppose Rs. 1,000 is given to you. How do you spend it?

Items	Amount (Rs.)
(a) Books	
(b) Clothes	
(c) Fast Food	
Total	1000

5.7   LET US SUM UP

There are four levels of measurements: nominal, ordinal, interval, and ratio.These constitute a hierarchy where the lower scale of measurement, nominal, has far fewer statistical applications than those further up this hierarchy of scales. Nominal scales yield data on categories; ordinal scales give sequences; interval scales begin to reveal the magnitude between points on the scale and ratio scales explain both order and the absolute distance between any two points on the scale.

The measurement scales, commonly used in marketing research, can be divided into two types; comparative and non-comparative scales. Comparative scales involve the respondent in signaling where there is a difference between two or more firms, brands, services, or other stimuli. The scales under this type are: (a) Paired Comparison, (b) Rank Order, (c) Constant Sum, and (d) Q-sort. Further, The non-comparative scales can be classified into: (a) Continuous Rating Scales and (b) Itemised Rating Scales. The Itemised Rating scales can further be classified into: (a) Likert Scale, (b) Semantic Differential Scale, and (c) Stapel Scale.

A number of scaling techniques are available for measurement of attitudes. There is no unique way that you can use to select a particular scaling technique for your research study. A number of issues, such as problem defintion and statistical analysis, choice between comparative and non-comparative scales, type of category lables, number of categories etc., discussed in this unit should be considered before you arrive at a particular scaling technique.

5.8   KEY WORDS

Comparative Scales : In comparative scaling, the respondent is asked to compare one object with another.

Constant Sum Scale : In this scale, the respondents are asked to allocate a constant sum of units such as points, rupees, or chips among a set of stimulus objects with respect to some criterion.

Continuous Rating Scales : Here the respondents rate the objects by placing a mark at the appropriate position on a continuous line that runs from one extreme of the criterion variable to the other.

Itemised Rating Scales : Itemised rating scale is a scale having numbers or brief descriptions associated with each category.

Interval Scale : In this scale, the numbers are used to rank attributes such that numerically equal distances on the scale represent equal distances in the characteristic being measured.

Likert Scale : With the Likert scale, the respondents indicate their own attitudes by checking how strongly they agree or disagree with carefully worded statements that range from very positive to very negative towards the attitudinal object.

Measurement : Measurement is the process of observing and recording the observations that are collected as part of research.

Non-comparative Scales : In non-comparative scaling, respondents need only evaluate a single object.

Nominal Scale : In this scale, the different scores on a measurement simply indicate different categories.

Ordinal Scale : In this scale, the items are ranked according to whether they have more or less of a characteristic.

Paired Comparison Scale : This is a comparative scaling technique in which a respondent is presented with two objects at a time and asked to select one object according to some criterion.

Q-Sort Scale : This is a comparative scale that uses a rank order procedure to sort objects based on similarity with respect to some criterion.

Rank Order Scale : In this scale, the respondents are presented with several items simultaneously and asked to order or rank them according to some criterion.

Ratio Scale : Ratio scales permit the researcher to compare both differences in scores and relative magnitude of scores.

Scaling : Scaling is the assignment of objects to numbers or semantics according to a rule.

Semantic Differential Scale : This is a seven point rating scale with end points associated with bipolar labels (such as good and bad, complex and simple) that have semantic meaning.

Staple Scale : The Staple scale places a single adjective as a substitute for the Semantic differential when it is difficult to create pairs of bipolar adjectives.

5.9   ANSWERS TO SELF ASSESSMENT EXERCISES

A. 1) Interval scale does not have a fixed (absolute) zero point whereas ratio scale has a fixed zero point that allows us to construct a meaningful ratio.

2) Knowing the level of measurement helps in interpreting the data and performing statistical analysis of the data.

3) In nominal scale the assigned numbers have no arithmetic properties and act only as labels. The only statistical operation that can be performed on nominal scales is frequency count.

4) a) Nominal Scale, b) Ratio Scale, c) Ordinal Scale, d) Interval Scale.

B. 1) Bias

2) 21

3) No. Some of the positive adjectives may be placed on the left side and some on the right side. This prevents the respondent with positive (negative) attitude from simply checking the left (right) side without reading the description of the words.

4) a) Itemised rating scale,   b) Paired comparison scale,   c) Constant sum scale.

5.10   TERMINAL QUESTIONS

1) Discuss briefly different issues you consider for selecting an appropriate scaling technique for measuring attitudes.

2) What are the different levels of measurement? Explain any two of them.

3) How do you select an appropriate scaling technique for a research study? Explain the issues involved in it.

4) Discuss briefly the issues involved in attitude measurement.

5) Differentiate between ranking scales and rating scales. Which one of these scales is better for measuring attitudes?

6) In what type of situation is the Q-sort technique more appropriate?

7) Name any four situations in commerce where you can use the Likert scale.

8) Construct a Rank Order Scale to measure toothpaste preferences. Discuss its advantages and disadvantages.

9) Construct a Semantic differential scale to measure the experiences of respondents in using Brand-X of shaving cream (assume that all the respondents use that brand).

> **Note:** These questions/exercises will help you to understand the unit better. Try to write answers for them. But do not submit your answers to the university for assessment. These are for your practice only.

5.11 FURTHER READING

The following text books may be used for more indepth study on the topics dealt with in this unit.

Aaker, David A. and George S. Day. (1983) *Marketing Research*, John Wiley, New York.

Bailey, Kenneth D. (1978) *Methods of Social Research*, The Free Press, New York.

Coombs, C.H.(1953) "Theory and Methods of Social Measurement", in *Research Methods in the Behavioral Sciences*, eds. Feslinger, L. and Ratz, D., Holt, Rinehart and Winston.

Donald S. Tull and Gerald S. Albaum. (1973) *Survey Research: A Decisional Approach*, Index Educational Publishers, New York.

Meister, David. (1985) *Behavioural Analysis and Measurement Methods*, John Wiley, New York.

Rodger, Lesile W. (1984) *Statistics for Marketing*, McGraw-Hill (UK), London.